Nov-Dec.
1961
v.65D
No.6

SCIENCE

Volume 65D
Number 6
Nov.–Dec. 1961

Radio Propagation

RECEIVED
DEC 4 - 1961
MIAMI UNIVERSITY
LIBRARY

U.S. DEPARTMENT
OF COMMERCE

SECTION D

JOURNAL OF RESEARCH

NATIONAL BUREAU OF STANDARDS

Journal of Research

of the

National Bureau of Standards

D. RADIO PROPAGATION

NOV.–DEC. • 1961

VOLUME • 65D

NUMBER • 6

Editor: James R. Wait
Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

Associate Editors: T. N. Gautier, J. W. Herbstreit, R. C. Kirby, C. G. Little, A. G. McNish, R. A. Helliwell, W. E. Gordon, A. D. Wheelon, S. Silver

IRE Advisors: A. H. Waynick, K. M. Siegel

Publication dates:
Jan. 15, Mar. 15, May 15, July 15, Sept. 15, Nov. 15, 1961

JOURNAL OF RESEARCH

The National Bureau of Standards Journal of Research reports research and development in the fields of activity shown at right. Also included from time to time are survey articles related to the Bureau's scientific and technical program.

The Journal is published in four separate sections as follows:

A. Physics and Chemistry

Contains papers of interest primarily to scientists working in these fields. Issued six times a year. Annual subscriptions: domestic, $4.00; foreign, $4.75.

B. Mathematics and Mathematical Physics

Presents studies and compilations designed mainly for the mathematician and the theoretical physicist. Issued quarterly. Annual subscription: domestic, $2.25; foreign, $2.75.

C. Engineering and Instrumentation

Reports research and development results of interest chiefly to the engineer and the applied scientist. Issued quarterly. Annual subscription: domestic, $2.25; foreign, $2.75.

D. Radio Propagation

Reports research in radio propagation' atmospheric physics, and communications. Issued six times a year. Annual subscription: domestic, $4.00; foreign, $4.75.

U.S. DEPARTMENT OF COMMERCE
Luther H. Hodges, Secretary

NATIONAL BUREAU OF STANDARDS
A. V. Astin, Director

Functions and Activities

The functions of the National Bureau of Standards are set forth in the Act of Congress, March 3, 1901, as amended. These include the development and maintenance of the national standards of measurement and the provision of means and methods for making measurements consistent with these standards; the determination of physical constants and properties of materials; the development of methods and instruments for testing materials, devices, and structures; advisory services to government agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; and the development of standard practices, codes, and specifications. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services, and various consultation and information services. The Bureau also serves as the Federal technical research center in a number of specialized fields. The scope of activities of the National Bureau of Standards is suggested in the following listing of the divisions and sections engaged in technical work.

Washington, D.C.

Electricity. Resistance and Reactance. Electrochemistry. Electrical Instruments. Magnetic Measurements. Dielectrics.

Metrology. Photometry and Colorimetry. Refractometry. Photographic Research. Length. Engineering Metrology. Mass and Scale. Volumetry and Densimetry.

Heat. Temperature Physics. Heat Measurement. Cryogenic Physics. Equation of State. Statistical Physics.

Radiation Physics. X-ray. Radioactivity. Radiation Theory. High Energy Radiation. Radiological Equipment. Nucleonic Instrumentation. Neutron Physics.

Analytical and Inorganic Chemistry. Pure Substances. Spectrochemistry. Solution Chemistry. Analytical Chemistry. Inorganic Chemistry.

Mechanics. Sound. Pressure and Vacuum. Fluid Mechanics. Engineering Mechanics. Rheology. Combustion Controls.

Organic and Fibrous Materials. Rubber. Textiles. Paper. Leather. Testing and Specifications. Polymer Structure. Plastics. Dental Research.

Metallurgy. Thermal Metallurgy. Chemical Metallurgy. Mechanical Metallurgy. Corrosion. Metal Physics.

Mineral Products. Engineering Ceramics. Glass. Refractories. Enameled Metals. Crystal Growth. Constitution and Microstructure.

Building Research. Structural Engineering. Fire Research. Mechanical Systems. Organic Building Materials. Codes and Safety Standards. Heat Transfer. Inorganic Building Materials.

Applied Mathematics. Numerical Analysis. Computation. Statistical Engineering. Mathematical Physics.

Data Processing Systems. Components and Techniques. Digital Circuitry. Digital Systems. Analog Systems. Applications Engineering.

Atomic Physics. Spectroscopy. Radiometry. Solid State Physics. Electron Physics. Atomic Physics.

Instrumentation. Engineering Electronics. Electron Devices. Electronic Instrumentation. Mechanical Instruments. Basic Instrumentation.

Physical Chemistry. Thermochemistry. Surface Chemistry. Organic Chemistry. Molecular Spectroscopy. Molecular Kinetics. Mass Spectrometry. Molecular Structure and Radiation Chemistry.

• Office of Weights and Measures

Boulder, Colo.

Cryogenic Engineering. Cryogenic Equipment. Cryogenic Processes. Properties of Materials. Gas Liquefaction.

Ionosphere Research and Propagation. Low Frequency and Very Low Frequency Research. Ionosphere Research. Prediction Services. Sun-Earth Relationships. Field Engineering. Radio Warning Services.

Radio Propagation Engineering. Data Reduction Instrumentation. Radio Noise. Tropospheric Measurements. Tropospheric Analysis. Propagation-Terrain Effects. Radio-Meteorology. Lower Atmospheric Physics.

Radio Standards. High-Frequency Electrical Standards. Radio Broadcast Service. Radio and Microwave Materials. Atomic Frequency and Time Interval Standards. Electronic Calibration Center. Millimeter-Wave Research. Microwave Circuit Standards.

Radio Systems. High Frequency and Very High Frequency Research. Modulation Research. Antenna Research. Navigation Systems. Space Telecommunications.

Upper Atmosphere and Space Physics. Upper Atmosphere and Plasma Physics. Ionosphere and Exosphere Scatter. Airglow and Aurora. Ionospheric Radio Astronomy.

Order all publications from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

Use of funds for printing this publication approved by the Director of the Bureau of the Budget (June 19, 1961).

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

The Solar Wind [1]

E. N. Parker

Contribution from Enrico Fermi Institute for Nuclear Studies and Department of Physics, University of Chicago, Chicago 37, Ill.

(Received June 19, 1961)

Hydrodynamic expansion of the solar corona is the basis for "solar corpuscular radiation." The quiet day coronal temperatures of 1 to 2×10^6 °K yield a solar wind of several hundred km/sec and a density of 5 to 50 particles/cm³ at the orbit of Earth. The solar wind draws out into space the lines of force of the general one-gauss solar field, to give an interplanetary field that is basically spiral in character with a density of the order of a few times 10^{-5} gauss at the orbit of Earth. The enhanced corona may sometimes have temperatures of 4×10^6 °K or more immediately following a large flare. Such explosive heating leads to a 1 to 2×10^3 km/sec blast wave into planetary space. The blast wave sweeps up the quiet-day solar wind ahead and kinks the quiet-day field, raising the density of the field to several times 10^{-4} gauss on some occasions. The galactic cosmic ray particles are swept back by the kink, producing a Forbush-type intensity decrease. Altogether, the blast wave and its magnetic field constitute the "enhanced solar corpuscular radiation."

It is shown that the alternative magnetic tongue model of Gold and others [1960] is untenable because the "tongue" would not reach the earth until many hours after the arrival of the blast wave.

1. Introduction

It was pointed out several years ago [Parker, 1958a] that the solar corona is too hot to be confined by the solar gravitational field. In place of the older view of the corona as an extended, but static, atmosphere it was shown that the corona must expand continually into space. A study of the hydrodynamic equations showed that the solar corona is steadly expanding with a velocity that is only a few kilometers per second in the lower corona but which becomes supersonic at large distances from the sun. The result is a continuing outward flow of coronal gas through interplanetary space, which we have called the *solar wind* to emphasize its hydrodynamic nature. The observed quiet-day coronal temperatures of 2×10^6 °K indicate that the quiet-day solar wind velocity is of the order of 150 to 500 km/sec. The density of the solar wind at the orbit of Earth is determined by the existing temperature observations to lie in the range 5 to 200 atoms/cm³. Thus the quiet-day solar wind represents a proton flux in the range 10^8 to 10^{10}/cm² sec at the orbit of Earth.

Some time before these calculations were made it had been pointed out by Biermann [1951, 1952, 1957] that the observed outward acceleration and the observed ionization and excitation of type I comet tails can be explained only as the result of continuing quiet-day solar corpuscular radiation. He originally estimated that the corpuscular flux was of the order of 10^{10} protons/cm² sec at the orbit of Earth, recently [Biermann, 1960] revising the estimate downward to 10^9 protons/cm² sec on the basis of laboratory measurements of the appropriate charge exchange cross sections.

Upon completion of the solar wind calculations it was at once obvious that Biermann's solar corpuscular radiation was in fact the hydrodynamic solar wind from the steadily expanding corona.

There are many other phenomena, such as the aurora, geomagnetic activity, etc., which are conventionally ascribed to solar corpuscular radiation. They too are the consequence of the solar wind. On an interplanetary scale the solar wind is hydrodynamic, but on a planetary scale its low density renders it corpuscular in many respects.

Since this first identification of solar corpuscular radiation with the expanding corona, we have explored a number of consequences of the solar wind [Parker, 1958 b, c; 1960, 1961 a, b]. It now appears that the solar wind is the dominant interplanetary dynamical force and is responsible for the interplanetary magnetic field configuration, the observed modulation of the galactic and solar cosmic ray intensity, the quiet day and the enhanced geomagnetic activity, etc.

Recent direct observation of the solar wind near Earth [Bridge, 1961] has determined its quiet-day velocity to lie in the range 250 to 400 km/sec, with a density of 10 to 20 protons/cm³. The total flux is of the order of 0.5×10^9 protons/cm² sec. An observation of the quiet-day solar wind in January and September of 1959 had been reported earlier by Shklovskii et al. [1960] to suggest a flux of 0.2×10^9 protons/cm² sec, but the rather large corrections that had to be applied to the data make it uncertain as to how meaningful the flux estimate really was.

[1] Paper presented at 1961 Spring meeting of the URSI, Washington, D.C.

2. The Expanding Corona

2.1. Quiet Sun

The simplest model of the quiet corona is one which assumes spherical symmetry about the center of the sun. Then if r is the radial expansion velocity as a function of distance from the center of the sun, the hydrodynamic equation for stationary expansion is

$$v\frac{dv}{dr}+\frac{1}{NM}\frac{dp}{dr}+\frac{GM\odot}{r^2}=0, \tag{1}$$

where N is the number of atoms per unit volume, G is the gravitational constant, and p is the hydrostatic pressure. Since the coronal gases are fully ionized and largely hydrogen, M is the mass of the hydrogen atom and the hydrostatic pressure is $p\cong 2NkT$. Conservation of mass requires that

$$Nvr^2=N_0v_0a^2 \tag{2}$$

where the subscript zero denotes the value at the reference level $r=a$. Choosing $a=10^6$ km, N_0 is of the order of $10^7/\text{cm}^3$. It will be sufficient for our purposes here to discuss the simple case of an isothermal corona. The more general case has been considered elsewhere [Parker, 1960].

For an isothermal corona, $T=T_0\cong 2\times 10^6$ °K, and (1) can be integrated to give

$$\frac{1}{2}(v^2-v_0^2)-\frac{2kT}{M}\ln\frac{v}{v_0}-\frac{4kT}{M}\ln\frac{r}{a}+\frac{GM\odot}{a}\left(1-\frac{a}{r}\right)=0 \tag{3}$$

upon using (2) to eliminate N. The result is a one parameter family of curves $v(r)$ for any given value of T_0, with v_0 as the parameter. The general form of the family is sketched in figure 1. The solution of physical interest is the solution starting from the origin and passing up through the critical point to supersonic velocity at infinity. The coordinates of the critical point are (v_c, r_c) where

$$\frac{1}{2}Mv_c^2=kT, GM\odot M/r_c=4kT. \tag{4}$$

The value of v_0 for this solution is the lower positive root of

$$\frac{Mv_0^2}{2kT}-\ln\frac{Mv_0^2}{2kT}=-3+\frac{GM\odot M}{akT}-4\ln\frac{GM\odot M}{4akT} \tag{5}$$

The coronal expansion will automatically drift into this solution whatever its initial expansion rate: The solutions with smaller v_0 go to $r=\infty$ on the lower branch of $v(r)$, which can be maintained only by a large inward pressure from $v=\infty$. In the absence of such an inward pressure the expansion will accelerate to the solut on through the critical point. There are no solutions above the solution through the critical point which start at the origin. Thus, the only available solution for expansion in the absence of an inward pressure at $r=\infty$ is the solution through the critical point, which starts with the velocity v_0 given by (5). It has been pointed out by Clauser [1960] that this expansion of the corona is analogous to the expansion of a gas through a Laval nozzle, with the gravitational field playing the role of the throat.

The expansion velocity as a function of radial distance is shown in figure 2 for a number of coronal temperatures. The observed coronal temperatures are of the order of 2×10^6 °K [Billings, 1959], but even low temperatures like 1×10^6 °K give solar wind velocities of the order of several hundred kilometers per second.

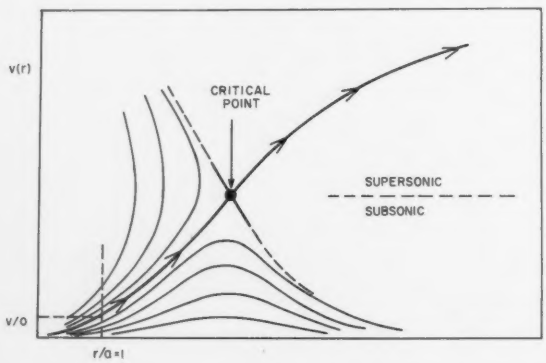

FIGURE 1. *The family of solutions of equation (3) for a given coronal temperature T_0.*

The solution of physical interest is the one passing through the critical point, given by eq (4). The critical point is analogous to the point of sonic transition in a Laval nozzle.

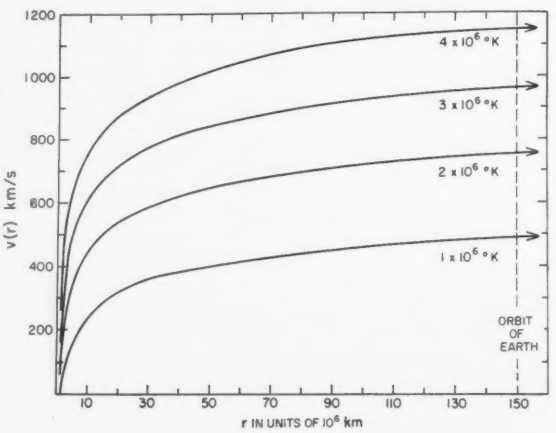

FIGURE 2. *The steady expansion velocity as a function of radial distance for a model isothermal corona.*

The two lower curves, for 1 and 2×10^6 °K, are applicable to the quiet corona. The upper curve, for 4×10^6 °K, gives an indication of the rate of expansion of the enhanced corona.

2.2. Active Sun

It is observed that following a large solar flare the coronal temperature over the general region of the flare may increase to 4×10^6 °K or more [Christiansen, Yabsley, and Mills, 1949; Christiansen and Warburton, 1953]. The result is a hydrodynamic explosion of the corona outward into space. The expansion velocity may be in excess of 1000 km/sec, as is evidence from figure 2. The asymptotic form of the resulting blast wave from the sun may be treated using the similarity transformations of the progressive wave [Courant and Friedrichs, 1949]. The idealization is made that the corona has spherical symmetry; i.e., the entire corona is heated to 4×10^6 °K, rather than the portion over the active region in which the flare occurred. The 300 km/sec velocity of the quiet day solar wind ahead of the blast wave is neglected, and the temperature of the quiet day wind is assumed to be so small that the 10^3 km/sec velocity of the blast wave represents a very high Mach number. The resulting blast wave profiles are shown in figure 3. The shock transition at the front is of the collisionless type and involves a density increase by a factor of four above the quiet-day value. The shape of the blast wave behind the shock transition depends upon how hard the enhanced corona pushes from behind. In the event that the corona should push so hard that the energy of the blast wave increases linearly with time after leaving the sun, then the relatively thin, high-density profile identified by $\lambda=1$ in figure 3 is the result. If the corona should not push at all once the blast wave is on its way, then a linear decrease of density behind the front is the result, identified by the sawtooth profile for $\lambda=3/2$ in figure 3. In either case an enhanced coronal temperature of 4×10^6 °K near the sun yields a blast wave with a velocity of $1\text{–}2\times10^3$ km/sec, and a density of the order of 10^2 cm^3 at the orbit of Earth [Parker, 1961, a and b]. Thus the blast wave represents the "enhanced solar corpuscular radiation" responsible for the geomagnetic storm, the Forbush decrease, etc.

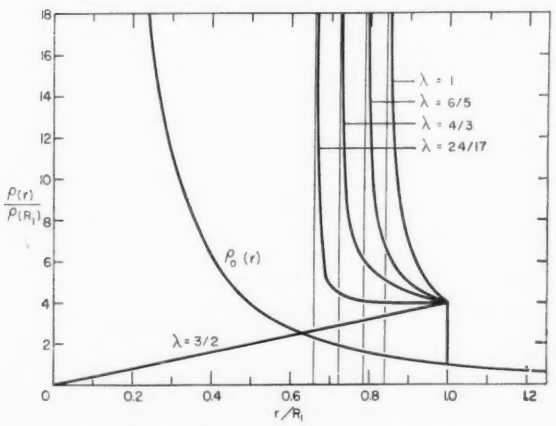

FIGURE 3. *Blast wave profiles from the active corona.*

The quiet day solar wind density is given by $\rho_0(r)$, and is essentially proportional to $1/r^2$. The shock transition at the head of the blast wave is at $r=R_1$, where the density jumps by a factor of four from the quiet day value. The parameter λ is an inverse measure of how strongly the enhanced corona is driving the blast wave from the rear: If the corona does not push at all on the blast wave, then $\lambda=3/2$ and the sawtooth profile is the result; if the corona pushes so hard that the energy of the blast wave increases linearly with time, then $\lambda=1$ and the blast wave lies between $r/R_1=0.84$ and $r/R_1=1.0$.

An important point to be noted is that the blast wave consists of gas which had already left the sun before the occurrence of the flare. It is material which has been swept up as the wave advances outward into space. The material in the corona at the time of the flare is behind the rear of the blast wave profile. This must be borne in mind in the next section when we discuss the interplanetary magnetic field configuration.

3. Interplanetary Magnetic Fields

3.1. Quiet Sun

The magnetic fields in interplanetary space are composed of the magnetic lines of force of the general solar field [Babcock and Babcock, 1955; Babcock, 1959] which are pulled outward from the sun by the expanding corona. The magnetic fields associated with active regions on the sun are usually considerably in excess of 1 gauss, so that they can to a large degree withstand the expansion of the local corona, remaining fixed over the active regions. The hydrostatic pressure in the solar corona is of such an order, 10^{-2} dynes/cm^2, as to suggest that only fields of the order of one gauss can be extended by the solar wind. The simplest model, then, is one in which a corona with spherical symmetry expands outward, carrying with it the dipole-like fields of the polar regions of the sun. Such an idealized picture ignores the many complications of the disordered one gauss fields which are undoubtedly carried outward from the equatorial regions. The idealization also ignores the many interesting complications that result from the fact that the actual corona expands more rapidly and densely in some directions than in others. However, the idealization of a radial solar wind with spherical symmetry in a dipole field serves to illustrate the general principles and to give a rough idea of the overall interplanetary magnetic configuration to be expected. When more comprehensive observations allow greater detail to be filled into the present simple theoretical models, the many irregularities and swirls which undoubtedly exist in the interplanetary field may be taken into account. Ignoring the irregularities now will give a conservative estimate of the cosmic ray effects to be discussed in the next chapter.

The lines of force of a general solar dipole field represent the two parameter family of curves

$$r-a\cong\frac{\gamma}{\Omega}(\phi-\phi_0),\ \theta=\theta_0$$

upon extension into space by the quiet day solar wind of velocity γ. The solar angular velocity Ω is

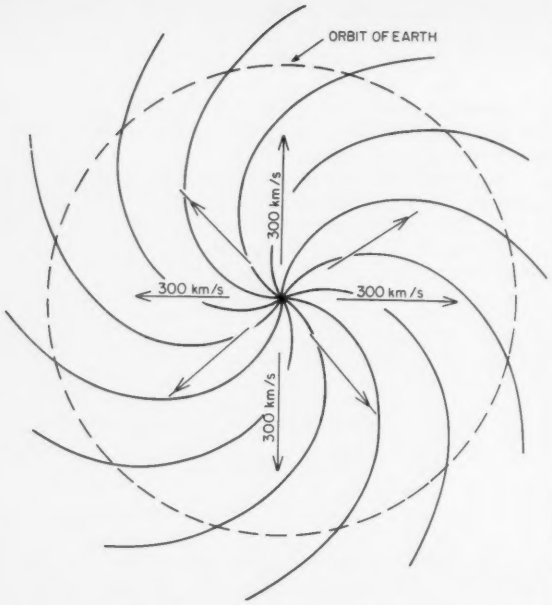

FIGURE 4. *Projection onto the equatorial plane of the lines of force of the solar fields extended by a quiet-day radial solar wind of 300 km/sec.*

approximately 2.6×10^{-6} radians/sec, and (θ_0,ϕ_0) is the position, in spherical polar coordinates, of the intersection of the line of force with the reference level $r=a$ in the lower corona. The radial component of the field is

$$B_r(r,\theta)=B_r(a,\theta)\left(\frac{a}{r}\right)^2,$$

the meridional component is zero, and the azimuthal component is

$$B_\phi(r,\theta)=B_r(a,\theta)r\Omega/v.$$

Projection of each line of force onto the solar equatorial plane is an Archimedes spiral, shown in figure 4 for $\gamma=300$ km/sec. It is readily seen that the field is principally radial inside the orbit of Earth and principally azimuthal beyond. One gauss at the solar photosphere yields about 2×10^{-5} gauss in the radial direction at a distance of one astronomical unit.

3.2. Active Sun

The quiet-day interplanetary field shown in figure 4 is deformed by the blast wave from the corona over a flare as shown in figure 5. Again the case $\lambda=1$ represents the situation when the enhanced corona is pushing hard on the rear of the blast wave, and $\lambda=3/2$ when the blast wave is coasting. In either case the field may reach 10×10^{-5} gauss at the front of the wave. When the corona is pushing hard on the rear of the blast wave the field toward the rear may easily be 40×10^{-5} gauss or more.

It has been suggested that the "enhanced solar corpuscular radiation," which we suggest is a hydrodynamic blast wave from the corona, sometimes draws out a loop of field from the active region, giving a magnetic tongue which is responsible for many of the effects associated with the enhanced corpuscular radiation (see discussion and references in Gold, 1960). Presumably the expanding interior of the magnetic tongue is partially shielded by the fields of the tongue from the galactic cosmic ray intensity, giving the Forbush-type cosmic ray decrease in coincidence with the geomagnetic storm resulting from the impact of the gas in the tongue against the geomagnetic field; once the tongue has engulfed Earth, it affords free access of solar protons from the sun to Earth; the tongue stores the solar protons following the emission from the sun, etc. In this way it has been argued that the tongue explains most of the observed cosmic ray phenomena associated with "enhanced solar corpuscular radiation."

We object to the tongue model of the interplanetary field on a number of points. First of all, the magnetic configuration of the blast wave fields shown in figure 5 possess the same properties as the tongue so far as modulating and channeling galactic cosmic rays and solar protons. Thus the tongue is not unique and seems an unnecessary complication. Second, and more serious, is the objection that the tongue cannot possibly arrive at Earth with the "enhanced solar corpuscular radiation." Whatever the nature of the outburst on the sun which leads to the enhanced radiation, the result is a blast wave which scoops up the interplanetary gas ahead to form a wave as shown in figure 3. The active coronal material is behind the blast wave, and hence so is the magnetic tongue which the coronal material is supposed to carry. Even in the extreme case that the enhanced corona is pushing so hard on the rear of the blast wave that $\lambda=1$ (which requires that the 10^6 °K coronal temperatures extend all the way to the rear of the wave) the tongue is 0.16 a.u. behind the front of the blast wave, which means a delay of the order of 6 hr in the arrival of the tongue. Yet it is clear in many cases that the cosmic ray modulation etc., begins with the first onset of geomagnetic activity. The tongue configuration is shown in figure 6 for the special case that $\lambda=3/2$. The tongue extends barely beyond the orbit of Mercury when the blast wave reaches the orbit of Earth.

4. Cosmic Ray Modulation

4.1. Quiet Sun

The general outward flow of nonradial magnetic fields from the sun tends to push back the galactic cosmic ray particles and thereby decrease the cosmic ray intensity in interplanetary space below the interstellar level. This reduction follows the general solar cycle and is usually referred to as the 11-year variation. There are many individual effects which probably contribute to the total reduction: the ideal-

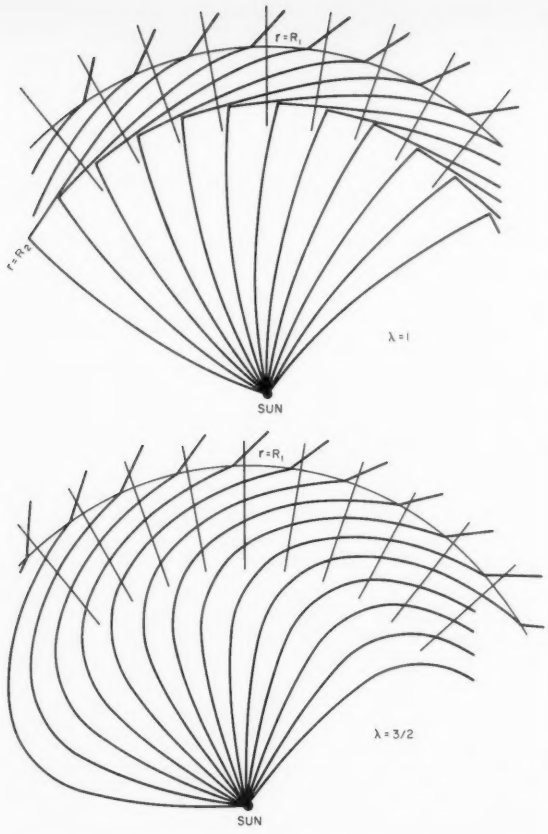

FIGURE 5. *Deformation of the quiet-day field shown in figure 4 by the freely coasting blast wave, $\lambda=3/2$, and by the driven blast wave $\lambda=1$.*

The shock transition is at $r=R_1$. The rear of the blast wave is at $r=R_2$.

ized quiet-day interplanetary field shown in figure 4 is, in effect, a giant corkscrew conveyor which revolves once every 27 days with the sun and tends to push outward the galactic particles sliding inward along the lines of force. There is reason to expect disordering in the interplanetary field, some examples of which have been cited elsewhere [Meyer, Parker, and Simpson, 1956; Parker, 1958c]; the disordering contributes to the trapping of solar particles and to the impediment of galactic cosmic ray particles diffusing inward from interstellar space.

These effects are large and contribute significantly to depression of the cosmic ray intensity. The depression is of the order of a factor of three at 1 Bev (see for instance Simpson, 1960) and probably more at lower energies. We may hope that one ' day interplanetary observation may decide what effects predominate.

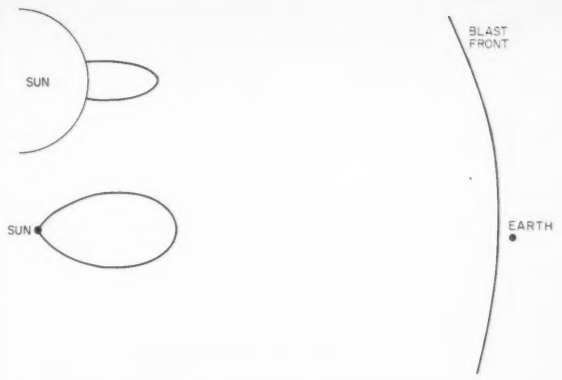

FIGURE 6. *A re-entrant loop of magnetic field in the solar corona before an outburst is shown, followed by its extension into space by the blast wave $\lambda=3/2$.*

The loop remains in association with the coronal gas with which it moves following the outburst. The blast wave ahead of the loop consists of interplanetary gas (the quiet-day solar wind) swept up to form the blast wave.

4.2. Active Sun

The magnetic configuration carried by the blast waves from the active corona, and shown in figure 5 for an idealized case, have the basic characteristic of a strong magnetic field (probably $10\text{–}50\times10^{-5}$ gauss) in the wave, with an essentially radial field connecting into the sun from behind. The strong fields result from the twisting and pinching together of the quiet-day lines of force. The resulting constriction in the lines of force represents a serious impediment to the passage of cosmic ray particles. Thus the blast waves will have a tendency to store energetic protons of solar origin behind them, and to push back galactic cosmic ray particles ahead of them, as required by observation (see for instance, Steljes et al., 1961). The Forbush decrease for the idealized model shown in figure 5 may be as large as 40 percent, and more if any complications are permitted in the idealized field configuration. The energy dependence of the cosmic ray decrease can be estimated if it is remembered that the blast waves are in fact not spherical but occupy a solid angle of the order of one steradian. The diffusion of cosmic rays from the sides of the region swept out by the blast waves gives an energy dependence to the fractional decrease of the cosmic ray intensity as shown in figure 7. The parameter ν in the figure is a measure of the angular width of the blast wave. The computed energy dependence is found to be not unlike the observed energy dependence of the Forbush decrease (see for instance MacDonald and Webber, 1960) viz, proportional to reciprocal rigidity, or flatter.

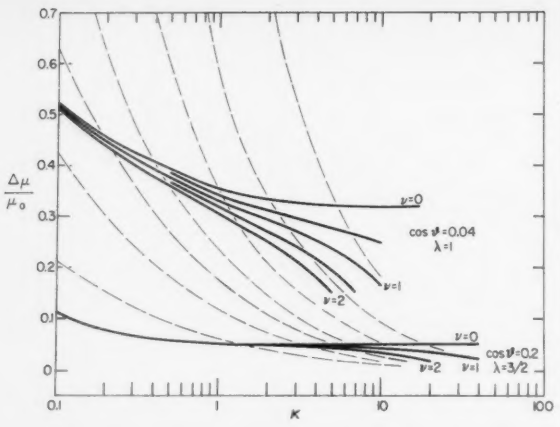

FIGURE 7. *The fractional depression $\Delta\mu/\mu$ of the cosmic ray intensity as a function of proton energy κ (in units of 931 Mev) for a freely coasting shock, $\lambda=3/2$, and for the hard driven shock, $\lambda=1$.*

The parameter ν is an inverse measure of the angular width of the blast wave as seen from the sun. The broken lines represent reciprocal rigidity, for purpose of comparison.

5. Summary

The purpose of the paper has been to describe in a semiquantitative way the solar wind model of interplanetary dynamical processes. The model affords a deductive approach to the problem of interplanetary plasmas, fields, and cosmic ray variations, starting with the hydrodynamic theory of the expanding solar corona. The observed coronal temperatures and their variations lead through such idealized models as we have discussed to the kind of "corpuscular" and cosmic ray effects that are known from observation. As direct observation of interplanetary conditions become more comprehensive with the progress of space experimentation, we look forward to filling in the many details which are missing from the present simple model.

6. References

Babcock, H. D., The sun's polar magnetic field, Astrophys. J. **130,** 364–365 (1959).

Babcock, H. W., and H. D. Babcock, The sun's magnetic field, 1952–1954, Astrophys, J. **121,** 349–366 (1955).

Biermann, L., Kometenschweife und solare Korpuskularstrahlung, Zeit. f. Astrophys. **29,** 274–286 (1951).

Biermann, L., Uber den Schweif des Kometen Halley im Jahre 1910, Zeit. f. Naturforsch. **7a,** 127–136 (1952).

Biermann, L., Solar corpuscular radiation and the interplanetary gas, Observatory **77,** 109–110 (1957).

Biermann, L., Paper presented at 4th Symposium on Cosmical Gas Dynamics, Varenna, Italy (August 1960).

Billings, D. E., Distribution of matter with temperature in the emission corona, Astrophys. J. **130,** 961–971 (1959).

Bridge, L., Paper presented at Spring Meeting of AGU, Washington, D.C. (April 1961).

Christiansen, W. N., and J. A. Warburton, The distribution of radio brightness over the solar disk at a wavelength of 21 centimetres, Australian J. Phys. **6,** 190–202 (1953).

Christiansen, W. N., D. E. Yabsley, and B. Y. Mills, Measurements of solar radiation at a wavelength of 50 centimetres during the eclipse of November 1, 1948, Australian J. Sci. Research **A2,** 506–523 (1949).

Clauser, F. H., Paper presented at 4th Symposium on Cosmical Gas Dynamics, Varenna, Italy (August 1960).

Courant, R., and K. O. Friedrichs, Supersonic flow and shock waves, pp. 419–421 (Interscience Publishers, New York, N.Y., 1948).

Gold, T. H., Energetic particle fluxes in the solar system and near the earth, Suppl. Astrophys. J. **4,** 406–416 (1960).

MacDonald, F. B., and W. R. Webber, Changes in the low-rigidity primary cosmic radiation during the large Forbush decrease of May 12, 1959, J. Geophys. Research **65,** 767–770 (1960).

Meyer, P., E. N. Parker, and J. A. Simpson, Solar cosmic rays of February 1956 and their propagation through interplanetary space, Phys. Rev. **104,** 768–783 (1956).

Parker, E. N., Dynamics of the interplanetary gas and magnetic fields, Astrophys. J. **128,** 664–676 (1958a).

Parker, E. N., Interaction of the solar wind with the geomagnetic field, Physics of Fluids **1,** 171–187 (1958b).

Parker, E. N., Cosmic-ray modulation by solar wind, Phys. Rev. **110,** 1445–1449 (1958c).

Parker, E. N., The hydrodynamic theory of solar corpuscular radiation and stellar winds, Astrophys. J. **132,** 821–866 (1960).

Parker, E. N., Sudden expansion of the corona following a large solar flare and the attendant magnetic field and cosmic-ray effects, Astrophys. J. **133,** 1014–1033 (1961a).

Parker, E. N., Sudden expansion of the corona following a large solar flare and the attendant magnetic field and cosmic-ray effects, Astrophys. J. **133,** 1014–1033 (1961b).

Shklovskii, I. S., V. I. Moroz, and V. G. Kurt, The nature of the earth's third radiation belt (in Russian), Astron. Zh. **37,** 931–934 (1960). For English translation see Soviet Astronomy **4,** No. 5, 871–873 (March–April 1961).

Simpson, J. A., Variations of solar origin in the primary cosmic radiation, Suppl. Astrophys. J. **4,** 378–405 (1960).

Steljes, J. F., H. Carmichael, and K. G. McCracken, Characteristics and fine structure of the large cosmic-ray fluctuations in November 1960, J. Geophys. Research **66,** 1363–1377 (1961).

(Paper 65D6–157)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Attenuation Coefficients for Propagation at Very Low Frequencies (VLF) During a Sudden Ionospheric Disturbance (SID) [1]

E. T. Pierce

Contribution from Stanford Research Institute, Menlo Park, Calif.

(Received May 11, 1961)

Attenuation coefficients for propagation under SID conditions are deduced, from records of atmospheric noise, for the frequency range of 3.5 to 50 kilocycles per second. These are compared with the values under normal daytime circumstances. It is shown that the advent of SID implies little change in attenuation between about 12 and 20 kilocycles per second; above this range there is a markedly decreased attenuation, while below 12 kilocycles per second there is an even more pronounced increase in the attenuation coefficients. The principle of an improved recorder of the changes in atmospheric noise associated with SID conditions is outlined; this recorder would discriminate between source effects independent of the SID and propagation influences solely attributable to the SID.

1. Introduction

The behavior, during propagation under normal daytime conditions, of radio waves of very low frequency (VLF), is now well established. It is known that the propagation is within the waveguide formed by the earth and the lower edge of the ionospheric D layer. Attenuation coefficients in the VLF range (3 to 30 kc/s) have been deduced using the mode theory [Wait, 1957], and have also been derived experimentally from observations of atmospherics [Chapman and Macario, 1956] and of longwave transmitters [Watt and Plush, 1959]. The agreement between experimental and theoretical results has been good, and the increasing recognition of such refinements as the influence of ground conductivity, the effect of the earth's magnetic field, and so on, is likely to lead to still better accord.

One influence that considerably affects daytime VLF propagation is the sudden ionospheric disturbance (SID). It is now generally agreed that a SID is caused by an increased flux of X-rays emitted by the sun during a solar flare [Friedman, 1959]. These X-rays enhance the ionization in the D layer. One aspect of this increased ionization is that an additional shallow bulge of electrons forms upon the underside of the normal D region. The exact shape of the bulge depends upon the spectral distribution of the X-rays in the incident flux [Pierce and Swift, 1960]; in all cases, however, the depression of the D layer will be greatest at the subsolar point, and there will be appreciable effects upon the D region within a radius of at least 6,000 km from the subsolar point.

As regards propagation at VLF the influence of a SID may be represented as a narrowing of the waveguide and, possibly, a change of the effective conductivity at the upper boundary of the guide. Consequences upon propagation may therefore be anticipated and these are indeed found. The main effects are sudden phase anomalies (SPA) in which a phase-change in the skywave from a manmade transmitter is detected, due to the depression of the D reflecting layer; sudden anomalies in field strength (SFA) which result from the changed phase relationship between the skywave and the groundwave, again from an artificial transmitter; and sudden enhancements of atmospherics (SEA) in which an increase of the noise due to distant thunderstorms is observed (typically at a frequency of 27 kc/s). SPA and SFA studies indicate that for a specific electron density an average SID lowers the D layer by some 5 km [Bracewell and Straker, 1949].

Although, as already mentioned, VLF attenuation coefficients have been widely obtained for normal conditions, the same information does not appear to have been derived under SID circumstances. The essential raw material, namely the changes observed in field strengths, has however been available for some time. It is the purpose of this note to attempt to obtain some indication of the VLF attenuation coefficients, during a SID, from data upon atmospherics.

2. Normal Daytime Attenuation Coefficients

Figure 1 depicts the relationship between the attenuation coefficient and frequency during normal daytime propagation. Two curves are shown. The first curve was deduced by Chapman and Macario [1956] from observations of atmospherics. The second is that derived by Watt and Plush [1959], largely on the basis of observations of manmade transmitters, and is described as being an average for a mixed path. There are divergencies between the two curves but these are not entirely unexpected. Even if D layer ionospheric conditions were uniform

[1] Paper presented at the 1961 Spring Meeting of the URSI, Washington, D.C.

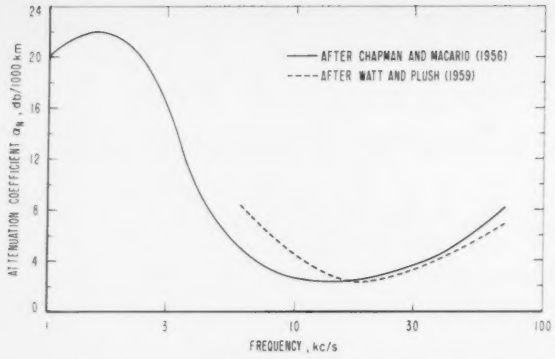

FIGURE 1. *Attenuation coefficients for normal daytime conditions.*

during daylight irrespective of solar altitude, such influences as ground conductivity and the direction of propagation with respect to the earth's magnetic field would still be important. These influences may lead to systematic deviations from "average" behavior. For instance, it is not unusual for the main sources of atmospherics received at a particular station to lie within a restricted geographical area; this implies that the attenuation coefficients, obtained from such observations of atmospherics, will be representative of a certain path, of its orientation with regard to the magnetic field and of the conductivity along its ground boundary, rather than of more general conditions.

3. Basic Data and Its Analysis

Wait [1958] has given the following expression for the vertical electric field E (in millivolts/meter) at a great circle distance d from a source radiating P kw:

$$E \simeq \frac{300}{h}\left\{\frac{P\lambda}{a \sin (d/a)}\right\}^{\frac{1}{2}} \exp(-\alpha d), \qquad (1)$$

where a is the radius of the earth, h the height of the ionospheric reflecting layer in kilometers, λ is the wavelength in kilometers, and α the attenuation factor. Equation (1) assumes that only a single mode is dominant, but this is a fair approximation up to 30 kc/s and for distances in the range 1,000 to 2,000 km, and almost completely true for d greater than 2,000 km; thus (1) is valid in the study of distant atmospherics. Equation (1), however, is likely to be increasingly incorrect as the frequency rises above 30 kc/s. Under normal conditions (suffix N) (1) may be written as

$$E_N = \frac{300}{h_N}\left\{\frac{P\lambda}{a \sin (d/a)}\right\}^{\frac{1}{2}} \exp(-\alpha_N d). \qquad (2)$$

If atmospheric noise is being received, then the advent of SID conditions (suffix S) will not cause P or d to change and

$$E_S = \frac{300}{h_S}\left\{\frac{P\lambda}{a \sin (d/a)}\right\}^{\frac{1}{2}} \exp(-\alpha_S d). \qquad (3)$$

Hence
$$h_N E_N / h_S E_S = \exp(\alpha_S - \alpha_N)d, \qquad (4)$$

it being assumed that E_N and E_S are compared at the same frequency. Equation (4) may be converted into the form

$$(\alpha_S - \alpha_N)d = 2\times 10^4 \log_{10}(h_N E_N / h_s E_s), \qquad (5)$$

in which the units for α_S and α_N are now the familiar db/1,000 km. Thus if records of E_N and E_S are available together with estimates for d, h_N, and h_S, $(\alpha_S - \alpha_N)$ can be deduced, and, knowing α_N, α_S obtained.

Four sets of data in the published literature give the ratio E_N/E_S at different frequencies and are therefore particularly suitable for the application of eq (5). These instances are in the papers by Gardner [1950], by Obayashi, Fujii, and Kidokoro [1959], and by Obayashi [1960]; this last paper contains two examples. Internal evidence within Gardner's paper suggests that 1,500 km is a good estimate for the distance of the noise sources; the reliability of this estimate is enhanced by the fact that Gardner worked in close collaboration with the Sferics Locating Network of the British Meteorological Office. It remains to determine h_N/h_S before $(\alpha_S - \alpha_N)$ can be calculated. h_N/h_S is not far removed from unity and the errors in its estimation are considerably less than those implicit in assigning a value to d. 70 km is a commonly accepted value for h_N, while h_S may be taken as (65 ± 5) km [Bracewell and Straker, 1949] the extremes corresponding to the slightest and to the most intense SID. Thus $h_N/h_S = 1.08$. Values of E_N/E_S were obtained from Gardner's paper and, using eq (5), $(\alpha_S - \alpha_N)$ was calculated over the frequency range 3.5 to 50 kc/s. The results are represented by curve (a) on figure 2.

Similar information to that of Gardner is given in the much later paper by Obayashi and coworkers [1959]. Here the frequency range is restricted to

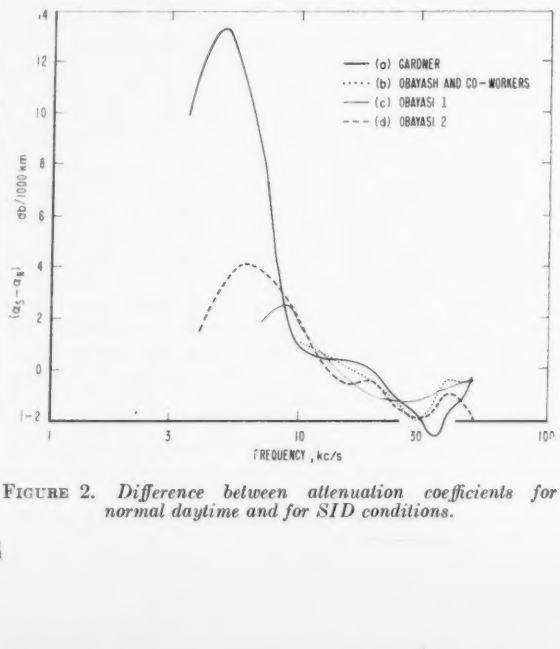

FIGURE 2. *Difference between attenuation coefficients for normal daytime and for SID conditions.*

10 to 50 kc/s. There is no direct evidence in the paper as to the distances of the sources of atmospheric noise. However, the SID concerned occurred between 0430 and 0500 U.T. on 16 August 1958. Atmospheric noise records for this day indicate no local sources, while sferics bearing observations covering the period of the SID show the activity to be to the east and to the southwest of Japan [IGY Data, Japan, 1959]. It seems plausible to associate these directions respectively with the main storm centers in the Americas and in SE Asia. The activity in the Americas will be dying with the advent of local night; that in SE Asia will be rapidly increasing the time being adjacent to local noon. In atmospheric noise measurements the nearest storms are the dominant factors [Horner, 1960]; thus the SE Asia center is likely to be more important than that in the Americas. Furthermore, it is the closer parts of the SE Asia center that will be most vital; this effect is enhanced by the influence of time since maximum thundery activity occurs in the local afternoon, and it is generally in the nearer parts of the SE Asia center to Japan, that, at the time of the SID, the local maximum is most closely approached. Considering all factors, it seems reasonable to identify the principal sources of the atmospherics as being within the nearest landmasses to the southeast of Japan, and to take d as being 2,000 km. Using $h_N/h_S=1.08$, the values of $(\alpha_S-\alpha_N)$ were calculated as before and are represented by curve (b) on figure 2.

Obayashi [1960] has given two further examples of enhancements of atmospheric noise which are suitable for analysis. These are for 29 March 1959 and 16 June 1959. The distance of the sources of noise for the first instance is estimated at 2,000 km, and Obayashi states that the storms were rather closer for the second enhancement; $d=1{,}500$ km would seem to be a reasonable mean estimate in this case. Using $h_N/h_S=1.08$ as before, the results obtained are shown in curves (c) and (d) on figure 2.

The agreement between the four curves in figure 2, for the range of 8 to 35 kc/s, is surprisingly good considering the speculative nature of the estimates for d in particular. Above 35 kc/s and below 8 kc/s, the signal strengths involved are low, and thus the results are more prone to errors. This effect may account for the spread between 35 and 50 kc/s, but seems quite inadequate to explain the discrepancies below 8 kc/s. Here the divergence between Gardner's and Obayashi's work is very considerable. The estimate for d is probably better for Gardner's observations than for Obayashi's; however, this factor is insufficient to enable the disagreements to be reconciled, and further experimental work is obviously desirable.

An intriguing feature of three of the curves on figure 2 is a maximum value and then a drop in $(\alpha_S-\alpha_N)$ as frequency decreases below 10 kc/s. If this trend is continued it is probable that at some lower frequency, possibly in the extremely low frequency (ELF) range (<3 kc/s), $(\alpha_S-\alpha_N)$ will become negative as it is at frequencies above about 15 kc/s. It might be predicted that a SID will produce an enhancement of atmospheric noise above 15 kc/s; a decrease at intermediate frequencies; and an enhancement again in the ELF band. This kind of predicted behavior, which would have analogies with the usual variation of α_N with frequency, could readily be verified by an extension of experimental observations to lower frequencies.

4. Discussion

Figure 3 is an attempt to summarize the preceding information. Curve (a) is an average representation for α_N derived from the two graphs of figure 1. Curve (b) is a composite graph for α_S; it was obtained by considering the four curves of figure 2, deducing weighted mean values for $(\alpha_N-\alpha_S)$, and using these, in conjunction with curve (a) of figure 3 to derive curve (b).

It is apparent from figure 3 that the frequency of 27 kc/s which has been established for many years [Bureau, 1937] as an indicator of enhancements in atmospherics is a good choice, since the minimum of α_S is close to 27 kc/s and the change between α_N and α_S is considerable. It is equally true that the differences between α_N and α_S are most marked for the lower frequencies, and the decrease in atmospherics over that frequency range is potentially a more sensitive indicator of the effects of a SID than is the enhancement at frequencies exceeding some 15 kc/s. This is particularly so if the normal storm centers are sufficiently close to the observing station for adequate daytime signal strength to be always present. An interesting feature of figure 3 is that the band (15 to 20 kc/s) occupied by the most important VLF transmitters is little affected in attenuation by the onset of SID conditions.

The normal atmospherics receiver on 27 kc/s which is usually intended as an indicator of the sudden enhancements associated with a SID, can also give increased signals produced by source effects and not by the propagation changes due to the SID. Three main causes may be assigned to such increases, namely, an augmentation of the

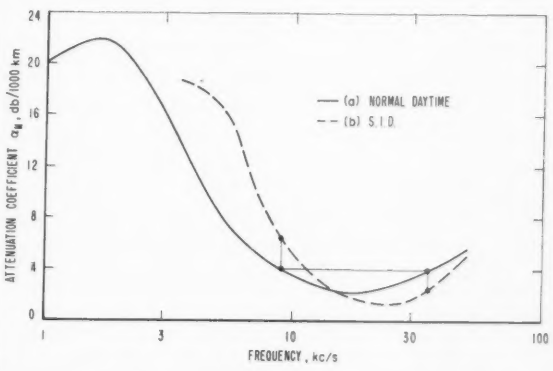

FIGURE 3. *Composite values for attenuation coefficients under normal daytime and under SID conditions.*

activity in an existing storm center; the development of a new storm center; and the movement of existing activity closer to the observing station. It seems possible, upon the basis of figure 3, to produce a reciprocal instrument which would not respond to the above source effects, but would only register a change in propagation conditions such as that produced by a SID. The principle would be to select two frequencies f_1 and f_2, one on each side of the normal minimum for α_N at around 18 kc/s, and for which the values of α_N are equal; as an example figure 3 indicates that 9 and 35 kc/s could be chosen for f_1 and f_2. Then if $E(f_1)$ and $E(f_2)$ are the measured fields at f_1 and f_2, the ratio $E(f_1)/E(f_2)$ should remain effectively constant irrespective of alterations in the strengths of the sources of atmospheric noise, or of changes in the distances of the sources, provided that these are not such as greatly to modify the validity of the basic one-mode assumption; this implies distances always greater than some 1,000 km. However if propagation conditions are affected as in a SID, the ratio $E(f_1)/E(f_2)$ will alter very considerably because of the changes in attenuation coefficient from α_N to α_S. The effect will be especially marked being reciprocal at the two frequencies, as is shown in figure 3 for the selected frequencies of 9 and 35 kc/s. Construction of a simple recorder presenting the ratio $E(f_1)/E(f_2)$ directly offers no great difficulty.

It is outside the scope of this paper to examine the interpretation of the α_S values of figure 3 upon the waveguide mode theory. However, there is no reason to believe that the values are incompatible with the theory, indeed the opposite is rather the case and the results are further evidence—although this is hardly necessary—of the general validity of the mode interpretation. For instance, it can readily be deduced from the work of Wait [1957] that a reduction in h from 70 to 60 km, without any considerable associated change in ionospheric conductivity, produces two graphs very similar to those depicted on figure 3. In each case, there is a reciprocal relative displacement of the two curves, pivoting around a common point.

References

Bracewell, R. N., and T. W. Straker, The study of solar flares by very long radio waves, Monthly Notices of Royal Astronomical Society, **109,** 28–45 (1949).

Bureau, R., Abnormalities of the ionosphere and bright solar eruptions, Nature **139,** 110–111 (1937).

Chapman, F. W., and R. C. Macario, Propagation of audio frequency radio waves, Nature **177,** 930–933 (1956).

Friedman, H., Rocket observations of the ionosphere, Proc. IRE **47,** 272–280 (1959).

Gardner, F. F., The use of atmospherics to study the propagation of very long radio waves, Phil. Mag. **41,** 1259–1269 (1950).

Horner, F., Properties of natural noise, Survey Paper before Commission IV, URSI XIIIth General Assembly, London, 1960.

IGY Data on Atmospherics, Whistlers, and Solar Radio Emissions, National Committee for the IGY, Japan, Vol. III for July–Dec. 1958, 1959.

Obayashi, T., Measured frequency spectra of very low frequency atmospherics, J. Research NBS **64D,** 41–48 (1960).

Obayashi, T., S. Fujii, and T. Kidokoro, An experimental proof of the mode theory of VLF ionospheric propagation, J. Geomagnetism and Geoelectricity, **10,** 47–55 (1959).

Pierce, E. T., and D. W. Swift, Study of spontaneous ionospheric disturbances, Research and Advanced Development Division, AVCO Corporation, Final Report for Geophysics Research Directorate, contract No. AF29(601)–1985 (1960).

Wait, J. R., The attenuation vs frequency characteristics of VLF radio waves, Proc. IRE **45,** 768–771 (1957).

Wait, J. R., A study of VLF field strength data: both old and new, Geofisica pura e applicata **41,** 73–85, (1958).

Watt, A. D., and R. W. Plush, Power requirements and choice of an optimum frequency for a world-wide standard-frequency broadcasting station, J. Reaserch NBS, **63D,** 35–44 (1959).

(Paper 65D6–158)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Dipole Radiation in a Conducting Half Space

R. K. Moore and W. E. Blair

Contribution from University of New Mexico, Albuquerque, N. Mex.

(Received August 31, 1960; as revised May 9, 1961)

The problem of communication between antennas, submerged in a conducting medium such as sea water, is analyzed in terms of a dipole radiating in a conducting half space separated by a plane boundary from a dielectric half space. The theory is discussed for both horizontal and vertical, electric and magnetic dipoles.

Expressions for the Hertzian potentials of the dipole in the conducting half space can be reduced to integrals obtained by Sommerfeld (for a dipole at the boundary) multiplied by an exponential depth attenuation factor. The Hertzian potentials are used to determine the electric and magnetic field components.

This analysis shows that the main path of communication between submerged antennas is composed of three parts as follows: (a) energy flow from the transmitting dipole directly to the surface of the sea, (b) creation of a wave that travels along the surface refracting back into the sea, (c) energy flow normal to the surface to the receiving dipole.

1. Introduction

Radiation from electric and magnetic dipoles in the air above a plane earth has been considered by many writers beginning with Sommerfeld [1909]. Sommerfeld's original paper treated a vertical monopole placed at the surface of a flat earth having arbitrary dielectric and conductive properties. The solution of the boundary-value problem was based on the evaluation of Fourier-Bessel integrals which were solutions to the wave equation. Weyl [1919] attacked the problem by resolving the dipole radiation into a spectrum of plane waves and evaluating the resulting integral without resorting to Sommerfeld's cylindrical wave formulation. Sommerfeld [1926] expanded his original work to take into account vertical and horizontal, electric and magnetic dipoles above a plane earth.

Sommerfeld's original paper contained an error in sign which was a subject of much later discussion [Norton, 1935]. Although Sommerfeld's work is followed here, the effect of the error in sign does not appear because of the assumption in this paper that the transmitting antenna is located in a highly conducting half space.

In general, papers published prior to 1940 on the subject of radiation of dipoles were concerned with dipoles in air or on the surface of the earth. Moreover, several papers have been written on this subject since 1940. Study of electromagnetic radiation in the sea began about the turn of the century and limited experimental work was done around the end of the first world war. Tai [1947] analyzed an electric dipole in an infinite, conducting medium and found extremely high dissipation in the vicinity of the dipole because of high conductive losses.

The first extensive theoretical treatment of electromagnetic radiation by vertical and horizontal, electric and magnetic dipoles immersed in a conducting half space appeared in a thesis by Moore [1951]. This paper describes a portion of that work. Wait [1952] [1957] considered an insulated magnetic dipole in an infinitely conducting medium, showing that the fields are independent of the characteristics of the insulation for an antenna diameter much less than the radiation wavelength in the conducting medium. Then Wait and Campbell [1953] and Wait [1959] analyzed a magnetic dipole of this type in a semi-infinite medium including special cases of frequency, antenna depth, and separation between antennas. Their model was a horizontal magnetic dipole (axis parallel to the surface of the sea), submerged in the sea (a conducting half space).

Analysis of the electric dipole was done by Lien and Wait [1953] in which the evaluation of the complex integrals was reduced to forms suitable for the numerical computation. Baños and Wesley [1953, 1954] carried out a mathematical analysis of the general case of a horizontal electric dipole in a conducting half space. The results in this paper agree with those of Baños

to the first order approximation. Kraichman [1960] experimentally verified Wait's and Baños' results in the intermediate distance region range. Also, the exponential increase of attenuation with depth was experimentally verified by Saran and Held [1960]. Anderson, in a thesis [1961], described the three-layered problem of earth-air-ionosphere for an electric dipole submerged in the sea. Further work by Wait [1960a, 1960b] has recently been carried out in an effort to unify the various approaches to the problem.

It is concluded from our analysis that the path of electromagnetic energy, between transmitting and receiving dipoles in the sea, is the following: (a) propagation from the transmitter directly to the surface, (b) propagation along the surface of the sea allowing refraction of the energy back into the sea, (c) leakage directly downward into the sea to the receiver.

A fundamental assumption in this analysis is that the displacement current (in the sea) is negligible compared to the conduction current. Since the conductivity of the sea is about 4 mhos/m and the frequencies of practical use in undersea communication are less than 50 kc/s, this is not at all a severe restriction for the case considered. Application to related problems depends upon the appropriateness of this assumption for each problem. The only other restriction is that the energy traveling directly through the sea between the transmitter and receiver is neglected. The ratio of the magnitude of the direct wave (through the sea) to the over-the-surface wave is of the order of $n\,e^{-(R-z)/\delta}$ where R, z, δ, and n, are respectively antenna separation, antenna depth, skin depth, and index of refraction of the sea with respect to the air. Thus for $R>>z$, the ratio is considerably less than unity.

Because of the "over-the-surface" mode of communication, the really important direction for transmitting dipole radiation is directly toward the surface of the sea. Since the radiation pattern of the vertical dipoles has a null in this direction, the vertical dipoles are not as effective in launching the required wave as are the horizontal dipoles. Electric and magnetic fields have been calculated for both the vertical and the horizontal dipoles, but the latter are of much greater importance in submarine communications.

2. Waves in Conducting Media

Discussions of the nature of electromagnetic radiation in a conducting medium and a conducting half space are presented by Stratton [1941] and Sunde [1949]. However, a brief review of the pertinent parts of the theory is included here for the purpose of defining the variables employed throughout this paper, as well as stating the fundamental conditions of the problem of dipole radiation in a conducting half space.

The Maxwellian curl equations for a conducting medium in which displacement current can be neglected are

$$\nabla\times\mathbf{E}=-j\omega\mu\mathbf{H} \tag{1}$$

$$\nabla\times\mathbf{H}=\sigma\mathbf{E} \tag{2}$$

where the time variation $e^{j\omega t}$ has been suppressed

$\mathbf{E}$ is the electric field vector (volts/meter)
$\mathbf{H}$ is the magnetic field vector (ampere-turns/meter)
ω is the angular frequency (radians/second)
μ is the magnetic permeability (henrys/meter)
σ is the conductivity (mhos/meter).

The diffusion equation for $\mathbf{E}$ is obtained by combining (1) and (2):

$$\nabla^2\mathbf{E}=j\omega\mu\sigma\mathbf{E}. \tag{3}$$

The solution of (3) for a plane wave traveling in the positive z-direction is:

$$\mathbf{E}=\mathbf{E}_0e^{-(1+j)z/\delta} \quad \text{with} \quad \mathbf{H}=e^{-j\pi/4}\sqrt{2}\,\frac{\mathbf{n}\times\mathbf{E}}{\mu\omega\delta}$$

where $\delta=\sqrt{2/\omega\mu\sigma}$ is the skin depth (meters)

$\mathbf{n}$ is the unit vector in the direction of propagation.

It is convenient in determining the fields of a dipole to determine first either the vector potential $\mathbf{A}$ or the Hertzian potential $\mathbf{\Pi}$. In this paper, the Hertzian potentials are used for both the electric and magnetic dipoles. The electric and magnetic field vectors are functions of the Hertzian potential as follows:

Electric dipole	*Magnetic dipole*	
$\mathbf{H}_1=\sigma\nabla\times\mathbf{\Pi}_1$	$\mathbf{E}_1=-j\omega\mu\nabla\times\mathbf{\Pi}_1^*$	(4)
$\mathbf{E}_1=\nabla\nabla\cdot\mathbf{\Pi}_1+k_1^2\mathbf{\Pi}_1$	$\mathbf{H}_1=\nabla\nabla\cdot\mathbf{\Pi}_1^*+k_1^2\mathbf{\Pi}_1^*$	(5)

where $\mathbf{\Pi}_1$, $\mathbf{\Pi}_1^*$ are the Hertzian potentials for the electric and magnetic dipoles respectively, and

$$k_1=\sqrt{-j\omega\mu\sigma} \qquad (\text{meters}^{-1}).$$

The subscript 1 denotes a conducting medium. Throughout this paper the expressions relating to the electric dipole are presented in the left column while those for the magnetic dipole are in the right column, side by side, for convenient comparison. The Hertzian potentials $\mathbf{\Pi}$ and $\mathbf{\Pi}^*$ (the Green's function for a dipole source) in the infinite, conducting medium are

$$\mathbf{\Pi}_1=\mathbf{p}_1\frac{e^{-jk_1R}}{R} \qquad \mathbf{\Pi}_1^*=\mathbf{p}_1^*\frac{e^{-jk_1R}}{R} \tag{6}$$

and

$$\mathbf{p}_1=\frac{\widehat{I\mathbf{l}}}{4\pi\sigma} \qquad \mathbf{p}_1^*=\frac{\widehat{NI\mathbf{S}}}{4\pi} \tag{7}$$

where R is the distance from the dipole to the observation point

$\widehat{I\mathbf{l}}=\lim_{\substack{I\to\infty\\ \mathbf{l}\to\mathbf{0}}}(I\mathbf{l})$, is the electric moment (ampere-meters)

$\widehat{NI\mathbf{S}}=\lim_{\substack{I\to\infty\\ \mathbf{S}\to\mathbf{0}}}(NI\mathbf{S})$ is the magnetic moment (ampere-meters ^{2})

I is the current (amperes).

$\mathbf{l}$ is the electric dipole length vector (meters).

N is the number of turns in the magnetic loop.

$\mathbf{S}$ is the loop area vector (meters ^{2}).

The corresponding statements for the electric and magnetic fields in an infinite, nonconducting medium are the following:

$\mathbf{H}_2=j\omega\epsilon_0\nabla\times\mathbf{\Pi}_2$	$\mathbf{E}_2=-j\omega\mu\nabla\times\mathbf{\Pi}_2^*$	(8)
$\mathbf{E}_2=\nabla\nabla\cdot\mathbf{\Pi}_2+k_2^2\mathbf{\Pi}_2$	$\mathbf{H}_2=\nabla\nabla\cdot\mathbf{\Pi}_2^*+k_2^2\mathbf{\Pi}_2^*$	(9)

where $k_2=2\pi/\lambda_0$, is the wave number (meters $^{-1}$)

λ_0 is free-space-wavelength (meters)

ϵ_0 is free space permittivity (farads/meter).

The Hertzian potentials in this case are

$$\boldsymbol{\Pi}_2=\mathbf{p}_2\frac{e^{-jk_2R}}{R} \qquad \boldsymbol{\Pi}_2=\mathbf{p}_2^*\frac{e^{-jk_2R}}{R} \tag{10}$$

$$\mathbf{p}_2=-j\frac{\widehat{\Pi}}{4\pi\omega\epsilon_0} \qquad \mathbf{p}_2^*=\frac{\widehat{NIS}}{4\pi}. \tag{11}$$

The above statement for the fields in an infinite medium must be modified for the case of a source in a conducting half space separated by a plane boundary from a nonconducting half space. For the modification, the source (transmitting dipole) is located in the conducting half space at coordinate position $(0, 0, z_t)$ from the boundary. Similarly, the point of observation (receiving dipole) is located at coordinate position (r, ϕ, z_r) (see fig. 1). The conducting half space (sea) is medium 1 (subscript 1), and the nonconducting half space (air) is medium 2 (subscript 2).

The Hertzian potentials in the following are normalized for convenience such that $|\mathbf{p}_1|=|\mathbf{p}_2|=|\mathbf{p}_1^*|=|\mathbf{p}_2^*|=1$. The Hertzian potentials in both media must satisfy the wave equation (condition I below) and the radiation condition at infinity (condition II) below:

(I)
$$\nabla^2\boldsymbol{\Pi}_1+k_1^2\boldsymbol{\Pi}_1=\mathbf{0} \qquad \nabla^2\boldsymbol{\Pi}_1^*+k_1^2\boldsymbol{\Pi}_1^*=\mathbf{0},\ z_r>0 \text{ (sea)}$$
$$\nabla^2\boldsymbol{\Pi}_2+k_2^2\boldsymbol{\Pi}_2=\mathbf{0} \qquad \nabla^2\boldsymbol{\Pi}_2^*+k_2^2\boldsymbol{\Pi}_2^*=\mathbf{0},\ z_r<0 \text{ (air)}$$

(II)
$$\lim_{R\to\infty}\boldsymbol{\Pi}_1=\mathbf{0} \qquad \lim_{R\to\infty}\boldsymbol{\Pi}_1^*=\mathbf{0} \qquad , z_r>0$$
$$\lim_{R\to\infty}\boldsymbol{\Pi}_2=\mathbf{0} \qquad \lim_{R\to\infty}\boldsymbol{\Pi}_2^*=\mathbf{0} \qquad , z_r<0$$

where $R=\sqrt{r^2+(z_t-z_r)^2}$.

Since the source is in the conducting half space, it is also necessary that Π_1 satisfy the radiation condition near the source:

(III)
$$\lim_{R\to 0}\Pi_1=\frac{e^{-jk_1R}}{R}=\lim_{R\to 0}\Pi_1^*.$$

In addition to the above three conditions, the electric and magnetic fields must satisfy the boundary conditions at the surface of the sea (the plane, $z=0$). Thus the boundary conditions for the components of the Hertzian potentials and their derivatives are the following:[1]

(IV)
$$-jg\Pi_{z1}=\Pi_{z2} \qquad \Pi_{z1}^*=\Pi_{z2}^*$$
$$-jg\Pi_{x1}=\Pi_{x2} \qquad -jg\Pi_{x1}^*=\Pi_{x2}^*$$
$$-jg\frac{\partial\Pi_{z1}}{\partial z}=\frac{\partial\Pi_{x2}}{\partial z} \qquad \frac{\partial\Pi_{z1}^*}{\partial z}=\frac{\partial\Pi_{x2}^*}{\partial z}$$
$$\frac{\partial\Pi_{x1}}{\partial x}+\frac{\partial\Pi_{z1}}{\partial z}=\frac{\partial\Pi_{x2}}{\partial x}+\frac{\partial\Pi_{z2}}{\partial z} \qquad \frac{\partial\Pi_{x1}^*}{\partial x}+\frac{\partial\Pi_{z1}^*}{\partial z}=\frac{\partial\Pi_{x2}^*}{\partial x}+\frac{\partial\Pi_{z2}^*}{\partial z}$$

where $g=\sigma/\omega\epsilon_0=|k_1^2/k_2^2|$. In condition (IV), it is assumed that the dipole is oriented either vertically in the z-direction, or horizontally in x-direction. (See fig. 1.) Sommerfeld [1949] has shown that only a Π_z component of the Hertzian potential is required to describe the fields of a vertical dipole. However, he has shown that both Π_x and Π_z components are required to describe the fields of a horizontal dipole (oriented in the x-direction).

---

[1] In condition (IV) the quantity $-jg$ is an approximation for the complex index of reflection squared, n^2. If the frequency or conductivity is such that $n^2\neq-jg$, then one can substitute the exact expression $n^2=-jg+\kappa$, where κ is the relative permittivity of the medium, throughout this paper.

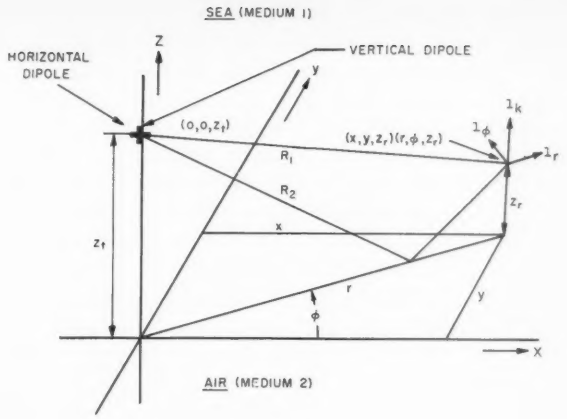

FIGURE 1. *Distances and coordinate system.*

3. Hertzian Potential Integrals

The Hertzian potential can be obtained in integral form for the four basic dipole configurations: vertical and horizontal, electric and magnetic dipoles. The method used here for obtaining the potential integrals was first used by Sommerfeld [1909, 1926]. The method is straightforward and available [Stratton, 1941; Sommerfeld 1949], so it will not be presented here. A general Hertzian potential for each of the four dipole configurations is obtained satisfying conditions (I) through (III). The general solution is then substituted into condition (IV) allowing the boundary conditions to determine the arbitrary coefficients of the general solution.

The Hertzian potentials for the observation point in the sea are presented in integral form for the four basic dipole configurations. The superscripts v and h denote vertical and horizontal dipole respectively. For $z_t>0$, $z_r>0$,

$$\Pi^v_{z1}=\mathscr{R}+I_{a1} \qquad \Pi^{*v}_{z1}=\mathscr{R}+I_{b1} \tag{12}$$

and

$$\Pi^h_{z1}=\mathscr{R}+I_{b1} \qquad \Pi^{*h}_{z1}=\mathscr{R}+I_{a1} \tag{13}$$

$$\Pi^h_{z1}=\cos\varphi\frac{\partial I_{c1}}{\partial r}=\Pi^{*h}_{z1} \tag{14}$$

where

$$\mathscr{R}=\frac{e^{-jk_1R_1}}{R_1}-\frac{e^{-jk_1R_2}}{R_2}$$

$$R_1=\sqrt{r^2+(z_r-z_t)^2} \qquad \text{(meters)}$$

$$R_2=\sqrt{r^2+(z_r+z_t)^2} \qquad \text{(meters)}$$

and where

$$I_{a1}=2\int_0^\infty \frac{e^{-F(z_r+z_t)}J_0(\xi r)\xi d\xi}{F-jgG} \tag{15a}$$

$$I_{b1}=2\int_0^\infty \frac{e^{-F(z_r+z_t)}J_0(\xi r)\xi d\xi}{F+G} \tag{15b}$$

$$I_{c1}=2(-jg-1)\int_0^\infty \frac{e^{-F(z_r+z_t)}J_0(\xi r)\xi d\xi}{(F+G)(F-jgG)} \tag{15c}$$

$$F=\sqrt{\xi^2-k_1^2} \qquad (\text{meters}^{-1})$$

$$G=\sqrt{\xi^2-k_2^2} \qquad (\text{meters}^{-1}).$$

The function $\mathscr{R}$ has two parts: exp $(-jk_1R_1)/R_1$ representing the primary source at the position (0, 0, z_t), and exp $(-jk_1R_2)/R_2$ representing a secondary source at the image position (0, 0, $-z_t$). In the case of the horizontal electric and the vertical magnetic antennas, the secondary source represents the image of the primary source. However, in the case of the vertical electric and the horizontal magnetic antennas, the secondary source represents an inverse image—that is, an image dipole of the opposite polarity. The primary source radiates over the direct path R_1, and the secondary source radiates over the reflected or image path R_2 (see fig. 1). Both paths R_1 and R_2 are entirely within the sea. Consequently, each of the Hertzian potentials in (12) and (13) is composed of three components: (a) a primary source function, (b) a secondary source function, and (c) an integral. It is shown later that the integral in every case represents the major contribution to the Hertzian potentials if $r>>(z_r+z_t)$. Therefore it is tacitly assumed that $\mathscr{R}$, composed of the primary and secondary sources, can be neglected. Consequently, (12) and (13) can be approximated for $z_r>0$, $z_t>0$, $r>>(z_r+z_t)$, as follows:

$$\Pi^e_{z1}\approx I_{a1} \qquad \Pi^{*h}_{z1}\approx I_{b1} \tag{16}$$

$$\Pi^h_{z1}\approx I_{b1} \qquad \Pi^{*h}_{z1}\approx I_{a1}. \tag{17}$$

I_{a1}, and I_{b1}, and I_{c1} are interdependent. It can be shown that [2] for $z_t>0$, $z_r>0$,

$$\frac{\partial I_{c1}}{\partial z_r}=I_{b1}+jgI_{a1} \tag{18a}$$

$$I_{c1}=\frac{\partial}{\partial z_r}\left[-\frac{2}{k_1^2}\frac{e^{-jk_1R_2}}{R_2}+\left(\frac{1}{k_2^2}+\frac{1}{k_1^2}\right)I_{a1}\right]\approx\frac{1}{k_2^2}\frac{\partial I_{a1}}{\partial z_r}. \tag{18b}$$

The Hertzian potentials for an observation point located in medium 2 are obtained in the same manner as were the Hertzian potentials for a point in medium 1. In the case of the potentials in medium 2, the function $\mathscr{R}$ is not involved, so the potentials can be written directly for $z_t>0$, $z_r<0$,

$$\Pi^e_{z2}=-jgI_{a2} \qquad \Pi^{*e}_{z2}=I_{b2} \tag{19}$$

and

$$\Pi^h_{z2}=-jgI_{b2} \qquad \Pi^{*h}_{z2}=-jgI_{a2} \tag{20}$$

$$\Pi^h_{z2}=-jg\cos\varphi\frac{\partial I_{c2}}{\partial r} \qquad \Pi^{*h}_{z2}=\cos\varphi\frac{\partial I_{c2}}{\partial r} \tag{21}$$

where

$$I_{a2}=2\int_0^\infty\frac{e^{-Fz_t+Gz_r}J_0(\xi r)\xi d\xi}{F-jgG} \tag{22a}$$

$$I_{b2}=2\int_0^\infty\frac{e^{-Fz_t+Gz_r}J_0(\xi r)\xi d\xi}{F+G} \tag{22b}$$

$$I_{c2}=2(-jg-1)\int_0^\infty\frac{e^{-Fz_t+Gz_r}J_0(\xi r)\xi d\xi}{(F+G)(G-jgG)} \tag{22c}$$

as in (18), the integrals I_{a2}, I_{b2}, and I_{c2} are also interdependent. For $z_t>0$, $z_r<0$,

$$\frac{\partial I_{c2}}{\partial z_r}=I_{b2}-I_{a2} \tag{23a}$$

$$I_{c2}=\frac{1}{k_2^2}\left(\frac{\partial I_{a2}}{\partial z_t}+\frac{\partial I_{a2}}{\partial z_r}\right). \tag{23b}$$

---

[2] (18) requires that the term $(-jg-1)$ in (15c) be kept in this form. However, for any practical calculation $(-jg-1)\approx-jg$ since $g>>1$ for the conductivity and frequencies of interest in this paper.

Now it is possible to write the electric and magnetic fields in medium 1 (sea) as a function of integrals I_{a1} and I_{b1}; and the fields in medium 2 (air) as functions of I_{a2} and I_{b2}. Since it is the purpose of this paper to determine the fields in the sea and at the surface of the sea, attention is focused on the integrals I_{a1} and I_{b1}. If the fields in the sea are determined, the boundary conditions can be used conveniently to determine the fields in air at the boundary ($z_r = 0-$). Consequently, *all* fields to be discussed in this paper can be expressed in terms of integrals I_{a1} and I_{b1}. It is convenient, however, to replace I_{a1} and I_{b1} with two new integrals I_a and I_b with a corresponding change of variables as follows:

$$\frac{c}{\omega} I_{a1}=I_a=2\int_0^\infty \frac{e^{-L\zeta}J_0(\rho\psi)\psi d\psi}{L-jgM} \tag{24}$$

$$\frac{c}{\omega} I_{b1}=I_b=2\int_0^\infty \frac{e^{-L\zeta}J_0(\rho\psi)\psi d\psi}{L+M} \tag{25}$$

where

$$L=\frac{c}{\omega} F=\sqrt{\psi^2+jg}$$

$$M=\frac{c}{\omega} G=\sqrt{\psi^2-1}$$

$$\psi=\frac{c}{\omega}\xi$$

$$\rho=\frac{\omega}{c} r=\frac{r}{\lambda_0/2\pi}$$

$$\zeta=\frac{\omega}{c} z=\frac{z}{\lambda_0/2\pi}$$

$z=z_r+z_t$ (meters)

c=speed of light in free space (meters/second)

λ_0=free-space wave length (meters).

The dimensionless distance factors ρ and ζ are used to express distances r and z in terms of (free-space-wavelength)/2π. Therefore the electric and magnetic field components in the sea, as functions of the integrals I_a and I_b are tabulated as follows:

Electric dipole	*Magnetic dipole*	
Vertical dipole		
$E_r=\frac{\omega^3}{c^3}\frac{\partial^2 I_a}{\partial\rho\partial\zeta}$	$H_r=\frac{\omega^3}{c^3}\frac{\partial^2 I_b}{\partial\rho\partial\zeta}$	(26)
$E_z=\frac{\omega^3}{c^3}\left(\frac{\partial^2 I_a}{\partial\zeta^2}-jgI_a\right)$	$H_z=\frac{\omega^3}{c^3}\left(\frac{\partial^2 I_b}{\partial\zeta^2}-jgI_b\right)$	(27)
$H_\varphi=-\frac{\omega^2}{c^2}\sigma\frac{\partial I_a}{\partial\rho}$	$E_\varphi=j\frac{\omega^3}{c^2}\mu\frac{\partial I_b}{\partial\rho}$	(28)
$E_\varphi=H_r=H_z=0$	$H_\varphi=E_r=E_z=0$	(29)
Horizontal dipole		
$H_r=\frac{\omega^2}{c^2}\sigma\sin\varphi\left(\frac{1}{\rho}\frac{\partial^2 I_a}{\partial\rho\partial\zeta}+\frac{\partial I_b}{\partial\zeta}\right)$	$E_r=-j\frac{\omega^3}{c^2}\mu\sin\varphi\left(\frac{1}{\rho}\frac{\partial^2 I_a}{\partial\rho\partial\zeta}+\frac{\partial I_a}{\partial\zeta}\right)$	(30)
$H_\varphi=\frac{\omega^2}{c^2}\sigma\cos\varphi\left(\frac{\partial^2 I_a}{\partial\rho^2\partial\zeta}+\frac{\partial I_b}{\partial\zeta}\right)$	$E_\varphi=-j\frac{\omega^3}{c^2}\mu\cos\varphi\left(\frac{\partial^2 I_a}{\partial\rho^2\partial\zeta}+\frac{\partial I_a}{\partial\zeta}\right)$	(31)

$$H_z=-\frac{\omega^2}{c^2}\sigma\sin\varphi\frac{\partial I_b}{\partial\rho} \qquad E_z=j\frac{\omega^3}{c^2}\mu\sin\varphi\frac{\partial I_a}{\partial\rho} \tag{32}$$

$$E_r=-\frac{\omega^3}{c^3}\cos\varphi\, jg\left(\frac{\partial^2 I_a}{\partial\rho^2}+I_b\right) \qquad H_r=-\frac{\omega^3}{c^3}\cos\varphi\, jg\left(\frac{\partial^2 I_a}{\partial\rho^2}+I_a\right) \tag{33}$$

$$E_\varphi=\frac{\omega^3}{c^3}\sin\varphi\, jg\left(\frac{1}{\rho}\frac{\partial I_a}{\partial\rho}+I_b\right) \qquad H_\varphi=\frac{\omega^3}{c^3}\sin\varphi\left(\frac{1}{\rho}\frac{\partial I_a}{\partial\rho}+I_a\right) \tag{34}$$

$$E_z=-\frac{\omega^3}{c^3}\cos\varphi\frac{\partial^2 I_a}{\partial\rho\partial\zeta} \qquad H_z=-\frac{\omega^3}{c^3}\cos\varphi\frac{\partial^2 I_a}{\partial\rho\partial\zeta}. \tag{35}$$

As already stated, the electric and magnetic fields in the air at the boundary can be determined from the fields in the sea and the boundary conditions. The fields are related through the boundary conditions as follows, for $z_r=0$:

$$H_{i2}=H_{i1} \qquad i=r,\varphi,z \tag{36a}$$

$$E_{i2}=E_{i1} \qquad i=r,\varphi \tag{36b}$$

$$E_{z2}=-jgE_{z1}. \tag{36c}$$

The boundary conditions (36) state that the magnetic fields and the tangential electric fields in air are equal, immediately adjacent to the boundary, to those in the sea, and that the vertical electric fields in the air are related to those in the sea by the proportionality constant $-jg$.

4. Evaluation of the Integrals

In the previous section the electric and magnetic fields in the sea and at the boundary were written as functions of the integrals I_a (24) and I_b (25). Apparently these integrals cannot be evaluated in closed form. Asymptotic expansions for I_a and I_b (as well as I_{a1}, (15a), I_{b1}, (15b), I_{a2} (22a), I_{b2} (22b)) and their appropriate derivatives can be determined by the method of critical points [Fredericks, 1953; Baños, 1953, 1954; Erdelyi, 1956].

However, it is possible to reduce the integrals I_a and I_b to those evaluated by Sommerfeld [1909, 1926, 1949] for the case in which the source and point of observation both lie on the boundary. Sommerfeld evaluated these integrals by transforming them into integrals along a contour in the complex plane. It is necessary to investigate the contours used to show the approximation possible for the case of the source and point of observation in the sea.

An integrand map is shown in figure 2. There are four branch points and two poles. The branch points are associated with both integrals, but I_b does not have a pole. The branch points arise because of the presence of the square roots in L and M. They occur where

$$L=\sqrt{\psi^2+jg}=0 \tag{37a}$$

$$M=\sqrt{\psi^2-1}=0. \tag{37b}$$

B.P. 1 and B.P. 3 are associated with L, and B.P. 2 and B.P. 4 are associated with M.

Coordinates of B.P. 1 and B.P. 3 are given by

$$\text{B.P. 1} \qquad \psi=\sqrt{g/2}(-1+j) \tag{38a}$$

$$\text{B.P. 3} \qquad \psi=\sqrt{g/2}(1-j). \tag{38b}$$

It can be seen that their distance from the origin is $\sqrt{g}$. For low frequencies and high conductivity, as with sea water, this is a rather large value (generally $> 10^3$ for communication in the sea).

In the form shown in (37b) it would appear that the branch points 2 and 4 would be on the real axis. It is obvious from the limits of the integrals that the path of integration must

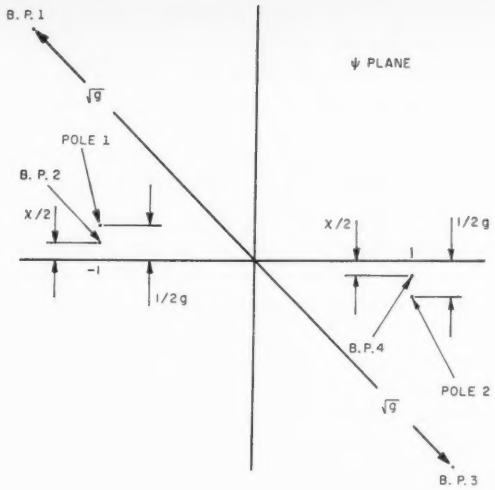

FIGURE 2. *Integrand map.*

lie along the real axis and this could cause certain complications. The reason this branch point appears to lie on the real axis is that medium 2 has been assumed to have zero conductivity. As a matter of fact, any practical medium, even air, has some conductivity and therefore the effective (complex) permittivity must contain a small imaginary component.

Thus, the complex permittivity, $\bar{\epsilon}$, is

$$\epsilon=\epsilon_0(1-j\chi) \tag{39}$$

where χ is a very small value associated with the conductivity of air. Utilizing the complex permittivity in (37b) we find the branch points are located by

$$\begin{aligned} &\text{B.P. 2} \quad \psi\approx -1+j\chi/2 \\ &\text{B.P. 4} \quad \psi\approx 1-j\chi/2. \end{aligned} \tag{40}$$

The poles occur where $(L-jgM)$ is zero. This occurs when

$$\psi^2=\frac{g^2}{1+g^2}\left(1-\frac{j}{g}\right). \tag{41}$$

For the case of large g (generally $> 10^6$) the location of the poles is given approximately by

$$\begin{aligned} &\text{Pole 1} \quad \psi\approx -1+j/2g \\ &\text{Pole 2} \quad \psi\approx 1-j/2g. \end{aligned} \tag{42}$$

The integrals as originally stated have a lower limit of zero and an infinite upper limit. To aid in evaluating them by contour methods it is convenient to write these integrals in forms such that the path of integration lies along the entire real axis. The principal consequence is the conversion of the Bessel functions of the first kind into Hankel functions of the first kind. Thus the integrals become

$$I_a=\int_{-\infty}^{\infty}\frac{e^{-L\zeta}}{L-jgM}H_0^{(1)}(\rho\psi)\psi d\psi \tag{43}$$

$$I_b=\int_{-\infty}^{\infty}\frac{e^{-L\zeta}}{L+M}H_0^{(1)}(\rho\psi)\psi d\psi. \tag{44}$$

It is therefore possible to close a contour by swinging a semicircle, whose radius is unbounded, from the positive real axis to the negative real axis through the upper half plane. Since there are two branch points in the upper half plane, one must form the contour such that the two associated branch cuts are not crossed. It is, of course, possible to choose convenient branch cuts, within limits.

Branch line 2 lies along the contour from $\psi=-1+j\chi/2$ to $\psi\to j\infty$ for which M is pure imaginary. It extends, in essence along the horizontal axis from -1 to 0 and then up along the imaginary axis. The contour of integration is shown in figure 3, where it can be seen that branch line 2 is the first branch cut encountered in swinging the infinite semicircle from positive real axis to negative real axis. The choice of values for the square root in M on the two Riemann surfaces separated by branch line 2 is such that for $z_r>0$, the exponent in the factor $e^{M\zeta}$ has a positive real part in the first quadrant of the ψ-plane and a negative real part in the second quadrant. (This aids in obtaining convergence of the integrals in the nonconducting medium (22) along branch line 1.) One observes as Sommerfeld did that in this case (high conductivity) Pole 1 and B.P. 2 are very close together so that the contours for I_a must pass around the branch point and the pole with the same loop. This problem has been treated in great detail by numerous authors and will not be elaborated here.

Branch line 1 extends from $\psi=\sqrt{g/2}\ (-1+j)$ to $\psi\to j\infty$. Along this line L has a large imaginary component and a small real component. The choice of Riemann surfaces is such that the real part of the exponent in the factor $e^{-L\zeta}$ is negative on the left side of the branch line, because this aids in exponential reduction of the integrand along branch line 2.

It can be shown readily that the integral along the infinite semicircle is zero. Following the terminology used by other writers in the past, the integral along branch line 2 is called $P+Q_2$ (except for I_b where it is just Q_2). The integral along branch line 1 is called Q_1. Thus for each of the integrals we may state that the value is given by

$$I=P+Q_1+Q_2 \tag{45}$$

where P is, of course, zero for I_b.

For a highly conducting medium, except very close to the source dipole, the contribution to the integral along branch line 1 is negligible. One can validate this statement by examining the asymptotic expansion of the zero order Hankel function of the first kind [Stratton, 1940] which is

$$H_0^{(1)}(\rho\psi)\sim\sqrt{\frac{2}{\pi\rho\psi}}\,e^{j(\rho\psi-\pi/4)}. \tag{46}$$

It is seen that the Hankel function vanishes exponentially as ψ increases along branch line 1. Since the smallest imaginary part of ψ associated with branch line 1 is $\sqrt{g/2}$, the contribution Q_1 is justifiably neglected. Consequently, it can be assumed that the significant contribution to the integral comes from the integration along branch line 2; that is

$$I\approx P+Q_2. \tag{47}$$

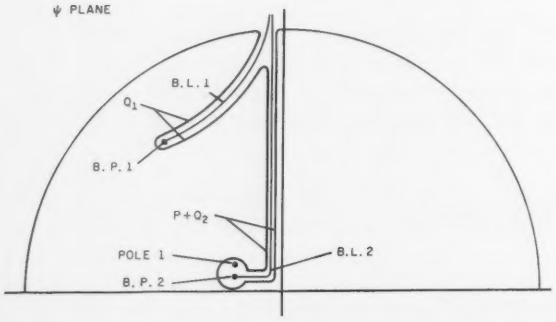

FIGURE 3. *Contour of integration.*

The *most important* approximation made in this treatment is the one allowing the removal of the exponential $e^{-L\zeta}$ from the integrand. To see that this is possible, it is necessary to express I, in the form of either (43) or (44), as the sum of two integrals for integration along branch line 2. The limits of the first integral are B.P. 2 and a point ψ_0 on branch line 2; and the limits of the second integral are ψ_0 and $j\infty$. The point ψ_0 is chosen such that $1<<|\psi_0|^2<<g$. For example, using $g=10^7$, one might choose $\psi_0=j10^3$. Consequently, one can approximate $L=\sqrt{\psi^2+jg}\approx\sqrt{jg}$ over the first integral (B.P. 2, ψ_0) with negligible error introduced in the evaluation of the first integral. Over the second integral ($\psi_0, j\infty$), the inequality

$$|e^{-\sqrt{jg}\zeta}|\leq|e^{-L\zeta}| \tag{48}$$

permits the following inequality to be written:

$$\left|\int_{\substack{\text{B.L.2}\\(\psi_0<\psi<j\infty)}}\frac{e^{-\sqrt{jg}\zeta}}{L+aM}H_0^{(1)}(\rho\psi)\psi d\psi\right|\leq\left|\int_{\substack{\text{B.L.2}\\(\psi_0<\psi<j\infty)}}\frac{e^{-L\zeta}}{L+aM}H_0^{(1)}(\rho\psi)\psi d\psi\right|=B \tag{49}$$

where $a=\begin{cases}-jg \text{ for } I_a\\ 1 \text{ for } I_b\end{cases}$

Using the approximation $M=\sqrt{\psi^2-1}\approx\psi$, over $(\psi_0, j\infty)$, and (46) with $10^{-2}<\rho<10^2, |\psi|\geq|\psi_0|\geq10^3$, the following upper bound is written for the inequality (49):

$$B\leq\left|\sqrt{\frac{2}{\pi\rho}}\int_{\psi_0}^{j\infty}\frac{e^{-L\zeta}e^{j\rho\psi-j\pi/4}2aj\psi^{3/2}d\psi}{jg+(1-a^2)\psi^2}\right|\leq\left|\sqrt{\frac{8}{\pi}}\frac{e^{-\sqrt{jg}\zeta+j\rho\psi_0}}{5jg\rho^{7/2}}(-j\rho^2\psi_0^2+2\rho\psi_0+2)\right| \tag{50}$$

It is shown later that a contribution whose magnitude is less than the upper bound (50) is negligible. Consequently, for the following approximation in I:

$$L\approx\sqrt{jg} \tag{51}$$

one observes that the error is negligible in the first integral and that the contribution of the second integral, where the approximation is relatively poor, is also negligible. Hence the approximation is justified. Using (51), the integrals I_a and I_b for the conductivity and frequency range considered in this treatment are written as follows:

$$I_a\approx P+Q_2\approx2e^{(-1-j)\sqrt{g/2}\zeta}\int_{\text{B. L. 2}}\frac{H_0^{(1)}(\rho\psi)\psi d\psi}{L-jgM} \tag{52}$$

and

$$I_b\approx Q_2\approx2e^{(-1-j)\sqrt{g/2}\zeta}\int_{\text{B. L. 2}}\frac{H_0^{(1)}(\rho\psi)\psi d\psi}{L+M}. \tag{53}$$

This approximation is important because it places outside the integral the factor associated with the distances of the dipole and of the point of observation from the boundary. Thus the remaining integral is simply that for observation at the surface of fields due to a source also located at the surface. This was the problem Sommerfeld [1909] treated for the vertical dipole, and the horizontal dipole problem has also been treated since then by many writers. *Thus the problem for fields due to a dipole in a conducting half space has been reduced to the known problem for fields due to a dipole at an interface.*

Both convergent and asymptotic series were developed by Sommerfeld and others for this integral. The expression for I_a was found [Moore, 1951] for $0<\rho<1$, to be

$$I_a = j2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{\rho g}\left\{1-\frac{1}{2}\frac{\rho}{g}+\frac{1}{2}\frac{j\rho}{1\cdot 3g^2}(1+j\rho)+\frac{1}{2}\frac{\rho}{1\cdot 5g^3}\left[1+j\rho+\frac{(j\rho)^{3/2}}{3}\right]+\cdots\right.$$

$$\left.+\frac{1}{2}\frac{j\pi}{\sqrt{jg}}H_0^{(1)}(-\rho)+\frac{1}{2}\frac{j\pi}{(jg)^{3/2}}H_1^{(1)}(-\rho)+\cdots\right\}. \tag{54}$$

This may be expressed as

$$I_a \approx j2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{\rho g}\,T. \tag{55}$$

For distances such that $10^{-2}<\rho<1$, conductivity $\sigma\approx 4$ mhos/m, and frequency $f<50$ kc/s, then $T\approx 1$. A more convenient form for I_a given by Sommerfeld [1949] is the following for $\rho>>1$:

$$I_a = j2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{g\rho}\left[1+\frac{j}{2}\sqrt{2\pi\eta}\;e^{-\eta/2}-\eta e^{-\eta/2}\int_0^{\sqrt{\eta}} e^{-u^2/2}du\right]. \tag{56}$$

An expression for I_b similar to that for I_a (54) was also obtained by Moore. However, I_b can be expressed in closed form [Wait, 1952] as follows:

$$I_b = 2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{(jg-1)\rho^3}\left[1+j\rho-(1+\sqrt{jg}\rho)e^{-(\sqrt{jg}-j)\rho}\right]. \tag{57}$$

For the practical purpose of submarine communications; that is, $\sigma\approx 4$ mhos/m, $f<50$ kc/s, it is sufficient to approximate I_a (54) or (56) and I_b (57) as follows:

$$I_a = j2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{g\rho}, \quad 10^{-2}<\rho<10^2, \quad \rho>>\zeta \tag{58}$$

and

$$I_b = -j2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{g\rho^3}(1+j\rho), \quad 10^{-2}<\rho\leq 1, \quad \rho>>\zeta \tag{59a}$$

or

$$I_b = 2\,\frac{e^{-j\rho-\sqrt{jg}\zeta}}{g\rho^2}\left(1-\frac{j}{\rho}\right), \quad 1\leq\rho<10^2, \quad \rho>>\zeta. \tag{59b}$$

Now that the integrals I_a and I_b have been evaluated, the derivatives of I_a and I_b involved in the field components ((26) through (36)) must be determined. It can be shown [Moore 1951] that the power series (54) can be differentiated, allowing the necessary derivatives to the same order as of (58) to exist. I_b (57) can be differentiated; consequently, the necessary derivatives of (59) also exist. Returning to the statement that $\mathscr{R}$ can be neglected in (16) and (17), one can use (16), (17), (24), (25), (58), and (59) to verify this approximation. Similarly, the approximation (51) can be justified by comparing the magnitude of the upper bound (50) to (58) or (59). For the case $g=10^7$, $\psi_0=j10^3$ and $10^{-2}<\rho<10^2$, the error is less than 1 percent.

5. Fields in a Conducting Half Space

After performing the indicated differentiations, substituting numerical values for permittivity, permeability, and velocity of propagation in free space, and inserting dipole moments, the resulting far field expressions (26) through (35) for $\rho>1$, are

$$E_r^{te} \approx -\sqrt{j}1.76\times 10^{-32}\,\frac{\omega^{7/2}Ile^{-z/\delta}}{\sigma^{3/2}\rho}\,T\left(1-\frac{j}{\rho}\right) \tag{60}$$

$$H_\varphi^{te} \approx -1.57\times 10^{-29}\,\frac{\omega^3 Ile^{-z/\delta}}{\sigma\rho}\,T\left(1-\frac{j}{\rho}\right) \tag{61}$$

$$E_\varphi^{tm} \approx 1.96\times 10^{-35}\,\frac{\omega^4 ISNe^{-z/\delta}}{\sigma\rho^2}\left(1-\frac{j2}{\rho}\right) \tag{62}$$

$$H_r^{vm} \approx \sqrt{-j}\, 1.76\times10^{-32} \frac{\omega^{7/2} ISNe^{-z/\delta}}{\sigma^{1/2}\rho^2}\left(1-\frac{j2}{\rho}\right) \tag{63}$$

$$H_r^{he} \approx -\sqrt{j}1.05\times10^{-23} \frac{\omega^{5/2} Ile^{-z/\delta}}{\sigma^{1/2}\rho^2}\left[\frac{1+T}{2}-\frac{jT}{2\rho}\right]\sin\varphi \tag{64}$$

$$H_\varphi^{he} \approx \sqrt{-j}\, 5.26\times10^{-24} \frac{\omega^{5/2} Ile^{-z/\delta}}{\sigma^{1/2}\rho} T\left[1-\frac{j}{T\rho}(2T-1)-\frac{2}{\rho^2}\right]\cos\varphi \tag{65}$$

$$E_r^{he} \approx -5.92\times10^{-27} \frac{\omega^3 Ile^{-z/\delta}}{\sigma\rho} T\left[1-\frac{j}{T\rho}(2T-1)-\frac{2}{\rho^2}\right]\cos\varphi \tag{66}$$

$$E_\varphi^{he} \approx j1.18\times10^{-26} \frac{\omega^3 Ile^{-z/\delta}}{\sigma\rho^2}\left[\frac{1+T}{2}-\frac{jT}{2\rho}\right]\sin\varphi \tag{67}$$

$$E_r^{hm} \approx -\sqrt{j}\, 6.64\times10^{-30} \frac{\omega^{7/2} ISNe^{-z/\delta}}{\sigma^{1/2}\rho} T\left(1-\frac{j}{\rho}-\frac{1}{\rho^2}\right)\sin\varphi \tag{68}$$

$$E_\varphi^{hm} \approx -\sqrt{-j}\, 1.33\times10^{-29} \frac{\omega^{7/2} ISNe^{-z/\delta}}{\sigma^{1/2}\rho^2} T\left(1-\frac{j}{\rho}\right)\cos\varphi \tag{69}$$

$$H_r^{hm} \approx j\, 1.17\times10^{-26} \frac{\omega^3 ISNe^{-z/\delta}}{\rho^2} T\left(1-\frac{j}{\rho}\right)\cos\varphi \tag{70}$$

$$H_\varphi^{hm} \approx -5.83\times10^{-27} \frac{\omega^3 ISNe^{-z/\delta}}{\rho} T\left(1-\frac{j}{\rho}-\frac{1}{\rho^2}\right)\sin\varphi. \tag{71}$$

Similarly, the resulting near field expressions ($\rho<1$) are

$$E_r^{ve} \approx \sqrt{-j}\, 1.76\times10^{-32} \frac{\omega^{7/2} Ile^{-z/\delta}}{\sigma^{3/2}\rho^2}(1+j\rho) \tag{72}$$

$$H_\varphi^{ve} \approx j1.57\times10^{-29} \frac{\omega^3 Ile^{-z/\delta}}{\sigma\rho^2}(1+j\rho) \tag{73}$$

$$E_\varphi^{vm} \approx -5.90\times10^{-35} \frac{\omega^4 ISNe^{-z/\delta}}{\sigma\rho^4}\left(1+j\rho-\frac{\rho^2}{3}\right) \tag{74}$$

$$H_r^{vm} \approx -\sqrt{-j}\, 5.26\times10^{-32} \frac{\omega^{7/2} ISNe^{-z/\delta}}{\sigma^{1/2}\rho^4}\left(1+j\rho-\frac{\rho^2}{3}\right) \tag{75}$$

$$H_r^{he} \approx \sqrt{-j}\, 1.05\times10^{-23} \frac{\omega^{5/2} Ile^{-z/\delta}}{\sigma^{1/2}\rho^3}(1+j\rho)\sin\varphi \tag{76}$$

$$H_\varphi^{he} \approx -\sqrt{-j}\, 5.26\times10^{-24} \frac{\omega^{5/2} Ile^{-z/\delta}}{\sigma^{1/2}\rho^3}(1+j\rho-\rho^2)\cos\varphi \tag{77}$$

$$E_r^{he} \approx 5.92\times10^{-27} \frac{\omega^3 Ile^{-z/\delta}}{\sigma\rho^3}(1+j\rho-\rho^2)\cos\varphi \tag{78}$$

$$E_\varphi^{he} \approx 1.18\times10^{-26} \frac{\omega^3 Ile^{-z/\delta}}{\sigma\rho^3}(1+j\rho)\sin\varphi \tag{79}$$

$$E_r^{hm} \approx \sqrt{j}\, 6.64\times10^{-30} \frac{\omega^{7/2} ISNe^{-z/\delta}}{\sigma^{1/2}\rho^3}(1+j\rho-\rho^2)\sin\varphi \tag{80}$$

$$E_\varphi^{hm} \approx -\sqrt{j}\, 1.33\times10^{-29} \frac{\omega^{7/2} ISNe^{-z/\delta}}{\sigma^{1/2}\rho^3}(1+j\rho)\cos\varphi \tag{81}$$

$$H_r^{hm} \approx 1.17\times10^{-26}\frac{\omega^3 ISNe^{-z/\delta}}{\rho^3}(1+j\rho)\cos\varphi \tag{82}$$

$$H_\varphi^{hm} \approx 5.84\times10^{-27}\frac{\omega^3 ISNe^{-z/\delta}}{\rho^3}(1+j\rho-\rho^2)\sin\varphi. \tag{83}$$

Here the superscripts *ve*, *vm*, *he*, and *hm* refer respectively to vertical-electric, vertical-magnetic, horizontal-electric, and horizontal-magnetic dipoles. T is determined from (54) and (55).

The fields are all expressed in terms of ρ, the radial distance in units of (free-space-wavelength)/2π. Since ρ is directly proportional to both ω and r, the frequency dependence of the result is not what it would be if ρ were independent of frequency. For example, at values of $10^{-2}<\rho<1$, several of the fields are independent of frequency except for the depth attenuation factor: E_r^{he}, E_φ^{he}, H_r^{hm}, and H_φ^{hm}.

The z-components of the fields in the conducting halfspace (sea) are all small (zero to a first-order approximation for the vertical electric dipole) compared with the horizontal components, and therefore they have not been included. On the other hand, the z-component of the electric field in the air, for the horizontal dipoles, is the predominant component. The expressions for the electric field vertical component in the nonconducting region at the surface can be obtained by applying the boundary condition (36c) for $\rho>1$; they are

$$E_{z2}^{he} \approx -\sqrt{-j}\ 1.98\times10^{-21}\frac{\omega^{5/2}Ile^{-z_1/\delta}}{\sigma^{1/2}\rho}T\left(1-\frac{j}{\rho}\right)\cos\varphi \tag{84}$$

$$E_{z2}^{hm} \approx 2.22\times10^{-24}\frac{\omega^3 ISNe^{-z_1/\delta}}{\rho}T\left(1-\frac{j}{\rho}\right)\sin\varphi. \tag{85}$$

Similarly, the results for $\rho<1$ are

$$E_{z2}^{he} \approx -\sqrt{j}\ 1.98\times10^{-21}\frac{\omega^{5/2}Ile^{-z_1/\delta}}{\sigma^{1/2}\rho^2}(1+j\rho)\cos\varphi \tag{86}$$

$$E_{z2}^{hm} \approx -j2.22\times10^{-24}\frac{\omega^3 ISNe^{-z_1/\delta}}{\rho^2}(1+j\rho)\sin\varphi. \tag{87}$$

These fields are given only at the surface as no effort has been made to evaluate the necessary field integrals at any distance away from the surface in the nonconducting medium. The horizontal electric antenna is the most practical for use in submarine communications. As an illustration of the manner in which the electric and magnetic fields vary as a function of distance from the source; the fields for the horizontal electric antenna have been plotted as a function of ρ in figures 4, 5, and 6, for typical values of frequency, conductivity, and practical antenna current, length, and depth.

6. Discussion and Conclusions

The form taken by the field equations indicates the mode of propagation that prevails as long as the distance from the boundary to both point of observation and source dipole is small compared with the horizontal separation. Each of the field expressions may be considered to consist of 3 parts: a multiplying factor which includes the dipole strength and parameters such as frequency and conductivity of the medium; an exponential attenuation factor whose exponent is the sum of the distance from the dipole to the surface and from the surface to the point of observation, expressed in units of skin depth; and a factor associated with variation in the radial (horizontal) direction. The latter factor is the same as that for surface fields of surface dipoles.

Thus it appears that propagation occurs as indicated in figure 7. The wave starts at the dipole, proceeds by the shortest path to the surface (the path of minimum attenuation), is refracted at the surface, travels along the surface as a wave in air, and comes to the point of observation by the path of least attenuation (straight down).

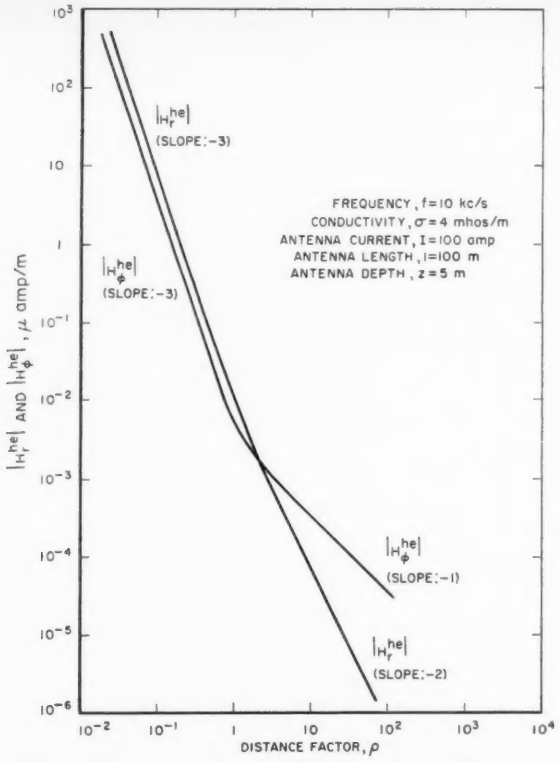

FIGURE 4. *Variation of the maximum absolute value of the magnetic field components in the sea for a submerged horizontal electric dipole with radial distance in units of free-space-wavelength/2π.*

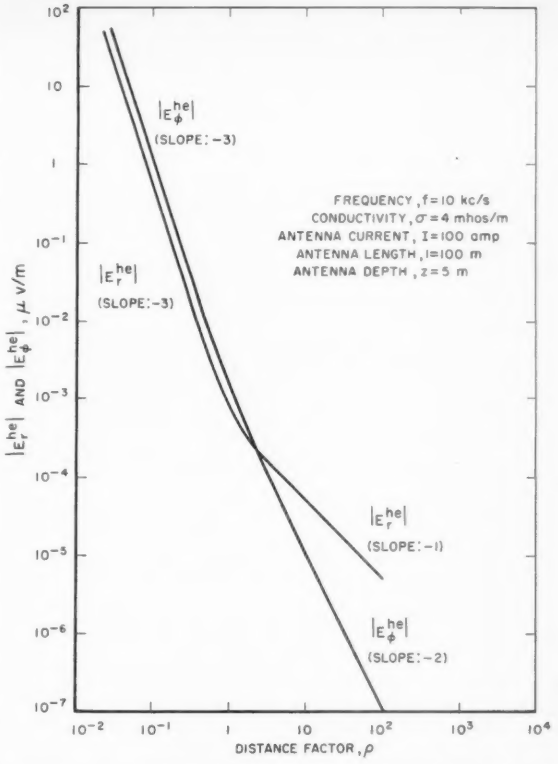

FIGURE 5. *Variation of the maximum absolute value of the electric field components in the sea for a submerged horizontal dipole with radial distance in units of free-space-wavelength/2π.*

It is possible to consider that an equivalent source on the surface sets up the wave in the nonconducting medium. This wave then travels as would be expected for such a source. Since the sea is not a perfect conductor, there is a slight tilt to the wave front, as indicated by the presence of both radial and axial components of the electric field; the radial (horizontal) component is much smaller than the axial (vertical) component. The Poynting vector for such a wave is nearly horizontal, but tilted slightly toward the conducting medium. The component of the Poynting vector associated with that part of the wave traveling into the conducting medium is small but finite. This wave which travels straight down is that which is seen by the observer in the conducting medium.

A similar situation prevails for waves in an imperfectly conducting waveguide. It is customary to calculate the losses in the wall of the waveguide by calculating the power flowing into the wall per unit area. In the waveguide this represents a leak of power from the desired wave. In this situation, where the observer is actually located in the conductor, the desired signal is the same as the "leak" for the waveguide.

Because of this mode of propagation, it is possible for radiation starting with the dipole in a conducting half space to be observed at much greater distances than would be possible if the wave were generated in an infinite homogeneous conducting medium. Thus, the exponential attenuation of 55 db per wavelength applies in the case of the conducting half space only to that part of the path from the dipole to the surface and from the surface to the point of observation. The rest of the path undergoes the same sort of attenuation that a wave generated by a dipole in air would encounter.

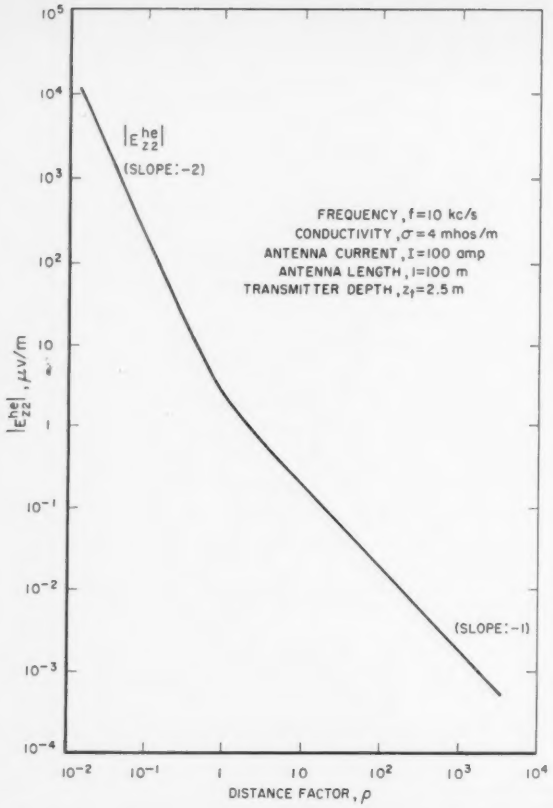

FIGURE 6. *Variation of the maximum absolute value of the normal electric field component in the air for a submerged horizontal electric dipole with radial distance in units of free-space-wavelength/2π.*

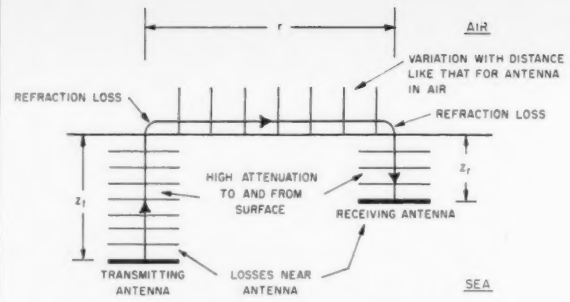

FIGURE 7. *Path of propagation.*

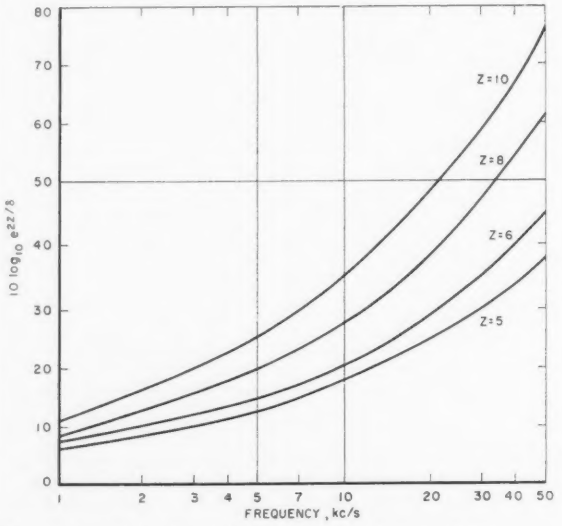

FIGURE 8. *Variation of depth attenuation factor with frequency.*

Although this development has applied to the case of a conducting half space and the expressions given are those associated with the approximation of the earth's surface by a plane, the fact that it is possible to separate out the horizontal attenuation of the wave in air means that it is also possible to apply the transmission loss that would be developed using spherical, rather than plane, earth geometry. In fact, it is also possible to use the attenuation calculated when the ionosphere is taken into account. Thus the results obtained here are considerably more general, within the limits of the approximations used, than would be indicated by the assumption of a plane earth.

It is interesting to note the magnitude of the attenuation of the wave in going through the conducting medium for the special case of sea water. This is shown for several depths as a function of frequency in figure 8. The depth indication is in meters. It is obvious that if the attenuation is to be kept low, the frequency must be kept low also.

It has been shown here that the fields created by a dipole in a conducting medium, which is near a plane boundary with a nonconducting medium, may be expressed in a form that may be considered as propagation from the dipole to the interface, travel through the nonconducting medium along the boundary, and propagation back into the conducting medium to the point of observation. Expressions have been presented for the fields of both vertical and horizontal, electric and magnetic dipoles. The fields of the horizontal dipoles are considerably stronger than those of the vertical dipoles, as would be expected because of the vertical nulls in the radiation from the vertical dipoles.

7. References

Anderson, W. L., The fields of electric dipoles in sea water—The water-air-ionosphere problem, D. Sc. Thesis, Univ. of New Mexico (1961).

Baños, A. Jr., and J. P. Wesley, The horizontal dipole in a conducting half-space, Univ. of Calif. Marine Physical Laboratory, Reports No. 53-33 (1953) and No. 54-33 (1954).

Erdelyi, A., Asymptotic expansions (Dover Publications, Inc., New York, N.Y., 1956).

Fredricks, K. O., Special topics of analysis, New York Univ. Inst. of Math. Sci. (1953).

Kraichman, M. B., Basic experimental studies of the magnetic field from electromagnetic sources immersed in a semi-infinite conducting medium, J. Research NBS **64D** (Radio Prop.), No. 1, 21–25 (Jan. 1960).

Lien, R. H., and J. R. Wait, Radiation from a horizontal dipole in a semi-infinite dissipative medium, J. Appl. Phys. **24,** 1–5 (Jan. 1953), 958–959 (July 1953).

Moore, R. K., The theory of radio communication between submerged submarines, Ph. D. Thesis, Cornell Univ. (1951).

Norton, K. A., Propagation of radio waves over a plane earth, Nature, **135,** 954–5 (June 1935).

Saran, G. S., and G. Held, Field strength measurements in fresh water, J. Research NBS **64D** (Radio Prop.), No. 5, 435–7 (Sept. 1960).

Sommerfeld, A., Über die Ausbreitung Elektromagnetischer Wellen über ein eben Erde, Ann. der Phys. Ser. 4, **28,** 665–736 (1909).

Sommerfeld, A., Über die Ausbreitung der Wellen in der drahtlosen Telegraphie, Ann. der Phys. Ser. 4, **81,** 1135–1153 (1926).

Sommerfeld, A., Partial differential equations in physics, Ch. VI (Academic Press, Inc. New York, N.Y., 1949).

Stratton, J. A., Electromagnetic theory (McGraw-Hill Book Co., Inc., New York, N.Y., 1940).

Sunde, E. D., Earth conduction effects in transmission systems, (D. Van Nostrand, Inc., New York, N.Y., 1949).

Tai, C. T., Radiation of a Hertzian dipole immersed in a conducting medium, Cruft Laboratory Report No. 21 (Oct. 1947).

Wait, J. R., The magnetic dipole antenna immersed in a conducting medium, Proc. IRE **40,** No. 10, 1244–5 (1952).

Wait, J. R., and L. L. Campbell, The fields of an oscillating magnetic dipole immersed in a semi-infinite conducting medium, Geophys. Research, **58,** No. 2, 167–178 (June 1953).

Wait, J. R., Insulated loop antenna immersed in a conducting medium, J. Research NBS **59,** No. 2, 133–137 (Aug. 1957).

Wait, J. R., Radiation from a small loop immersed in a semi-infinite conducting medium, Can. J. Phys. **37,** 672–674 (1959).

Wait, J. R., Propagation of electromagnetic pulses in a homogeneous earth, Appl. Sci. Research, Sec. B, **8,** 213–253 (1960a).

Wait, J. R., The electromagnetic fields of a horizontal dipole in the presence of a conducting half-space, Sci. Rep. No. 11, Contract CSO and A 58-40, Air Force Cambridge Research Labs., 1960b (to be published in Can. J. Phys. (July 1961).

Weyl, H., Ausbreitung elektromagnetischer Wellen über einem ebener Leiter, Ann. der Phys. Ser. 4, **60,** 481–500 (1919).

(Paper 65D6–159)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Reliability of Atmospheric Radio Noise Predictions[1]

John R. Herman

Contribution from AVCO Corporation, Research and Advanced Development Division, Geophysics Section, 201 Lowell Street, Wilmington, Mass.

(Received May 15, 1961)

Measured radio noise values are compared with the corresponding International Radio Consultative Committee [C.C.I.R., 1957] predicted values at four noise measuring stations. Five frequencies between 0.013 and 10.0 Mc/s are considered. The stations selected for this study include Balboa, Panama, near two major radio noise centers, and Byrd Station, Antarctic, remote from atmospheric radio noise sources. It is found that the predicted and measured noise levels are in good agreement except at some places and times, where large discrepancies occur. Most of the disagreements are found at places where the predictions are based on extrapolations of data measured at other stations. Reasons for the disagreements are discussed.

1. Introduction

Because the success or failure of communications by radio waves depends upon the signal-to-noise ratio at the receiver, it has become necessary to obtain some knowledge of the noise levels to be expected. Atmospherically-generated radio noise levels vary with time of day, season, and geographical location. Prediction of the variations on a worldwide basis have been published in Radio Propagation Unit (R.P.U.) Technical Report No. 5 [1945, 1947] and National Bureau of Standards Circular 462 [1948] in terms of the minimum field strengths required to maintain intelligible voice communications ninety percent of the time in the presence of noise. Subsequently, the prediction curves were revised in NBS Circular 557 [Crichlow, 1955] to show expected median levels of radio noise power during four-hour time blocks for each season. As more recent data accumulated it became apparent that large discrepancies existed between the measured and predicted values at some locations and times, so that an entirely new set of predictions was prepared and published in International Radio Consultative Committee (C.C.I.R.) Report No. 65 [1957]. The C.C.I.R. predictions closely follow those in Circular 557, but are based on a wider selection of compatible data, and should thus be more useful than the previous publications.

Cooperative measurements of received noise power at 16 widely separated geographical locations (fig. 1) are currently being conducted on a routine basis under the direction of the National Bureau of Standards Central Radio Propagation Laboratory. Since the issuance of the C.C.I.R. predictions, several additional noise measuring stations have been established and approximately 5 years of data have been collected. Thus it is now possible to check the accuracy of the predictions during a time of declining sunspot activity (1960), and at geographical locations where the noise level could only be interpolated when the predictions were established.

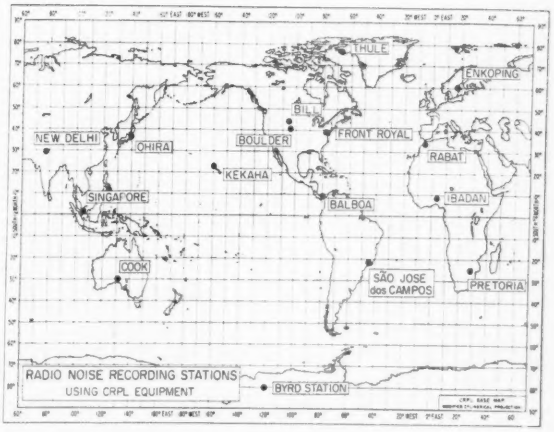

FIGURE 1. *Map showing location of noise measuring stations (reproduced from NBS Tech. Note* 18).

In this paper comparisons are made of the predicted and measured noise values at the four stations listed in table 1 for the frequencies (in Mc/s) shown. The letter X indicates the frequencies analyzed at each station. These particular stations were selected because their noise grades range from high to very low. That is, the noise grade at Balboa is high because it is near two major noise centers in Central and South America, while that at Byrd Sta-

TABLE 1

Station and location	Frequency (Mc/s)							
	0.013	0.051	0.246	0.495	0.545	2.5	5.0	10.0
Balboa, Panama, 9° N, 80° W	-----	X	X	X	-----	X	X	X
Byrd Sta., Antarctica, 80° S, 120° W	-----	X	-----	-----	X	X	X	X
Enkoping, Sweden, 60° N, 17° E	-----	X	-----	-----	X	X	X	X
Kekaha, Hawaii, 22° N, 160° W	X	-----	-----	X	-----	X	X	X

[1] This work was supported by the U.S. Air Force Cambridge Research Laboratories under Contract No. AF19(604)–4092.

tion is very low because it is remote from all atmospheric noise sources. The noise grades at Kekaha and Enkoping are between these two extremes; that at Kekaha being greater than at Enkoping.

2. Noise Level Predictions

The C.C.I.R. predictions are presented as worldwide noise level contour maps for 1.0 Mc/s in 4-hr time blocks for each season. Curves are given to determine the noise levels at other frequencies between 10 kc/s and about 30 Mc/s. The magnitude of the noise is in terms of a basic parameter representing the median value of atmospheric noise power passed by a narrow bandwidth, expressed for a 4-hour time block. The seasons of the year for the present purpose are defined as in table 2.

The basic parameter, F_a [Crichlow, 1955], used to describe the noise is defined as the noise power available from a short, vertical, grounded, loss-free antenna in decibels above kTB, where kTB is the terminal noise power in watts, available from a passive resistance (dummy antenna). Here,

k = Boltzmann's constant = 1.38×10^{-23} joules per degree Kelvin.

T = absolute room temperature = 288° K.

B = effective noise bandwidth in cycles per second (c/s).

TABLE 2. *Seasons of the Year (in accordance with C.C.I.R. Report No. 65 and NBS Technical Note 18)*

Months	Season	
	Northern Hemisphere	Southern Hemisphere
December, January, February	Winter	Summer.
March, April, May	Spring	Autumn.
June, July, August	Summer	Winter.
September, October, November	Autumn	Spring.

2.1. Prediction Curves for Each Station

The expected atmospheric noise level at each of the four noise measuring stations listed in table 1 are shown in figures 2 to 5 by solid lines for each season of the year and each frequency. The corresponding measured noise values are shown in the figures by dashed lines, and the estimated manmade noise levels at the stations are depicted as broken lines. (All times referred to in this paper are local times.) More about the measured values is given in the next section.

The expected atmospheric noise levels shown in figures 2 to 5 were prepared from the noise grade maps given in the C.C.I.R. report. Because the noise maps are for 1.0 Mc/s only, the predicted values of F_a for the frequencies considered here were extrapolated using the curves of noise power as a function of frequency given in figures 21 and 22 of that report. In a few cases, notably at Marie Byrd Station and at Kekaha for some time blocks, the interpolation between noise isopleths on the maps is rather broad, but the chosen values should be correct within one or 2 db.

The estimated manmade noise levels shown in figures 2 to 5 were also obtained from the C.C.I.R. report. The values represent what should be expected at receiving sites located several miles from populated areas and at least a quarter of a mile from elevated power lines and electrical machinery.

3. Noise Level Measurements

3.1. Seasonal Time-Block Measured Values

The noise power is recorded for 15 min per hr for each frequency, and each sampling is assumed to represent the noise level for the full hour. From these data the month-hour medians, F_{am}, are obtained in terms of decibels above kTB and are compiled in the NBS Technical Note 18 Series [Crichlow, 1959 a and b]. Also given are the seasonal 4-hr time block values, averaged over each 4-hr period for the 3 months of the season in accordance with C.C.I.R. recommendations. The tabulations include the upper and lower decile values of F_a to give an indication of the extent of the day to day variations from the median, F_{am}. The appropriate seasonal time-block noise power averages as reported by Crichlow [1959b] are plotted as dashed lines in figures 2 to 5.

Since the noise-measuring equipment approximates, insofar as practicable, the idealization of the system for which the predictions were generated, and corrections have been made in the measured data for antenna losses and the like, the measured and predicted values of noise power given in figures 2 to 5 can be compared directly.

3.2. Comparison of Measured and Predicted Seasonal Time-Block Values

As is evident from the curves of figures 2 to 5, a great number of discrepancies exist between the predicted and measured values of noise power. In several cases there is very good agreement, but in others the predicted noise level is as much as 40 db lower or 24 db higher than that actually measured. However, it should be noted that in most of the disagreements the predicted noise level is on the low side, and a predicted value 24 db too high occurs only once (winter 1600–2000 hour time block at Byrd Station on 0.051 Mc/s) in the data analyzed here. Furthermore, most of the too-high predicted values occur at Byrd Station and Enkoping, where the prediction curves are based on scanty data or extrapolation techniques.

Measured values that are much higher than the corresponding predicted values can be attributed to a number of causes. One of the most important causes is manmade noise near the receiving station. This can arise from automobile ignition systems, power lines, and electrical machinery, which is propagated primarily over power lines or by ground waves. Since manmade noise is random it is present on all the measured frequencies. Another important contribution to the measured noise level might arise from local station interference, and this would show up on only one frequency. Atmospherics generated by local thunderstorm activity would

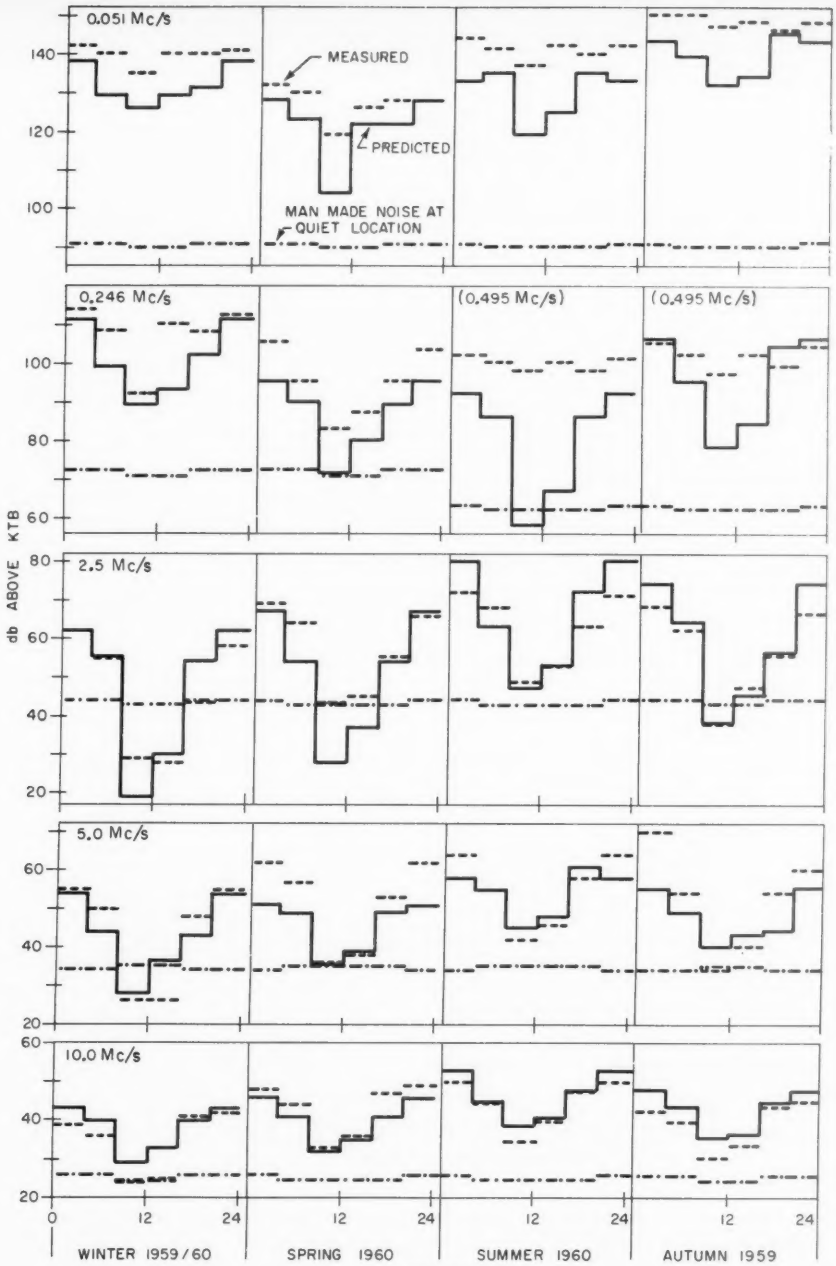

FIGURE 2. *Predicted and measured noise power at Balboa, Panama.*
Prediction curves scaled from C.C.I.R. Report No. 65.

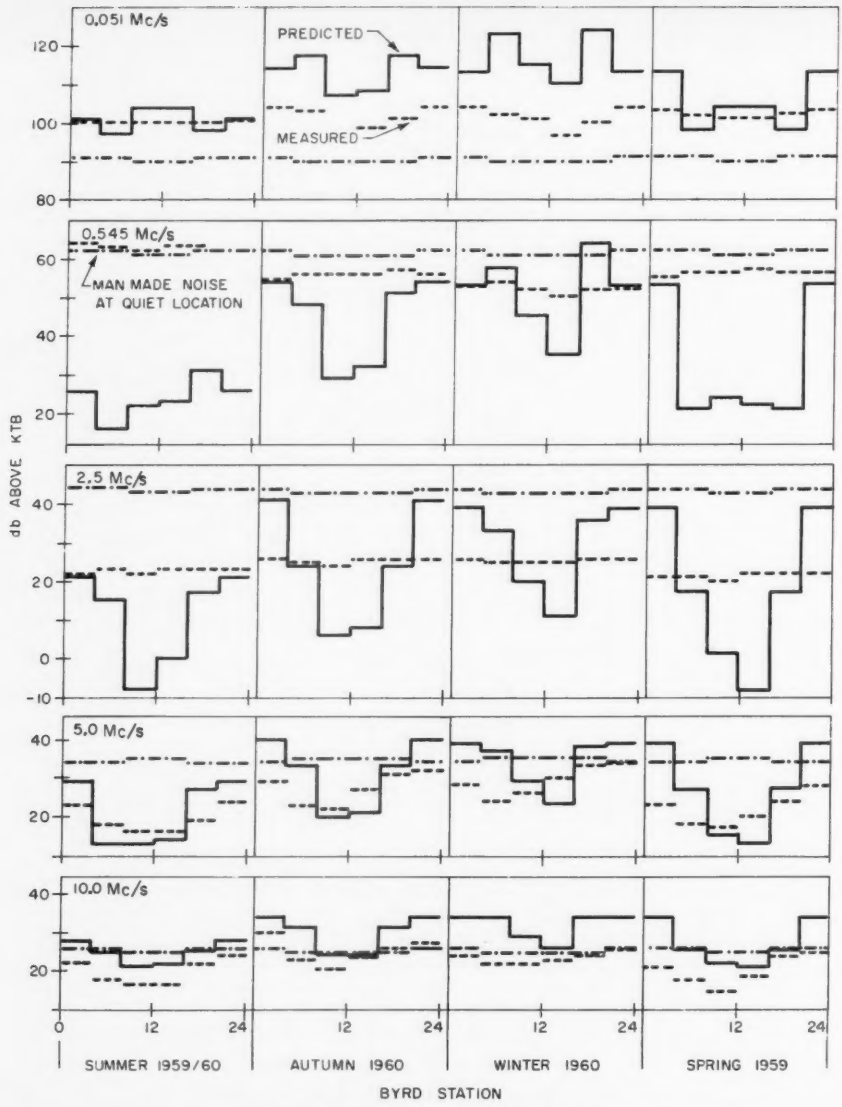

FIGURE 3. *Predicted and measured noise power at Byrd Station, Antarctica.*
Prediction curves scaled from C.C.I.R. Report No. 65.

greatly increase the received noise power, but this has been taken into account in the prediction curves from a knowledge of worldwide thunderstorm activity. Thus, unless the locale of a receiving station suffered a higher-than-usual number of thunderstorms spaced over a 3-month period, local atmospheric noise would not be the cause of a measured noise value much higher than that predicted in a seasonal time block.

The prediction curves were issued for atmospheric noise levels primarily, while the noise data [Crichlow, 1959 a and b] include radio noise of all types; manmade, cosmic, and atmospheric; which happened to be incident on the antenna at the time of measurement. Data which obviously contained manmade noise, interfering signals or equipment trouble have been omitted from the compilations, but these effects have not been eliminated in all cases [Crichlow, 1959 b]. Thus, the high noise level measured at Balboa on 0.495 Mc/s in the spring and summer (fig. 2) probably contains a strong contribution from a local interfering signal. Were it manmade noise or local atmospherics, the corresponding data on 2.5, 5.0, and 10.0 Mc/s would show its effects. The higher-than-predicted

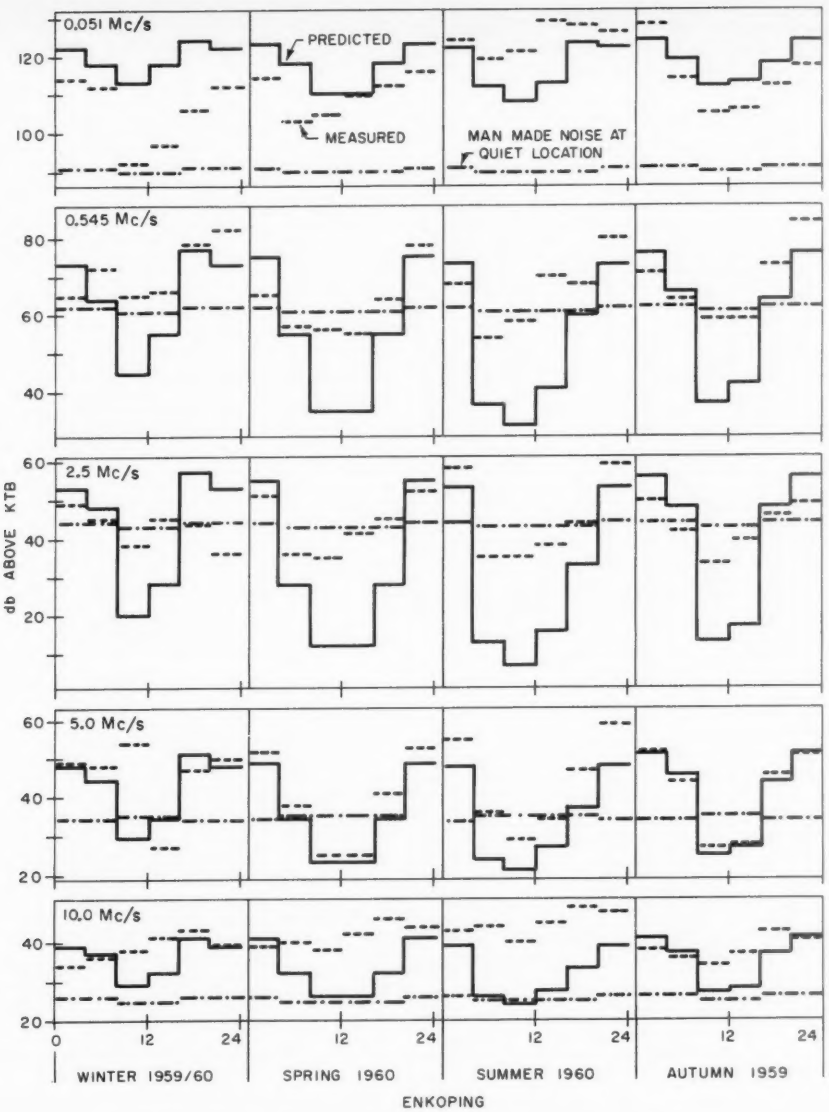

FIGURE 4. *Predicted and measured noise power at Enkoping, Sweden.*
Prediction curves scaled from C.C.I.R. Report No. 65.

noise values measured at Balboa on 5.0 Mc/s during the night hours (2000–0400) of spring, summer, and autumn also apparently contain an interfering signal. In this case, however, the signal appears to have been ionospherically propagated from a distant transmitter, because it disappears during the daytime when ionospheric absorption is effective.

The data from Byrd Station (fig. 3) are of particular interest because this station is located in a very low noise region where no previous noise measurments had ever been made. The C.C.I.R. report suggests caution in the use of predicted values which fall below the estimated manmade noise levels for a quiet receiving site, and this is just the case for the frequencies of 0.545, 2.5, and 5.0 Mc/s at Byrd Station. Since thunderstorm activity is practically nonexistant at Byrd Station, the atmospheric radio noise must be propagated in from distant sources, such as the major noise centers in equatorial Africa, South America, and the East Indies. Other contributions to the measured noise level would, of course, arise from local manmade sources and interfering

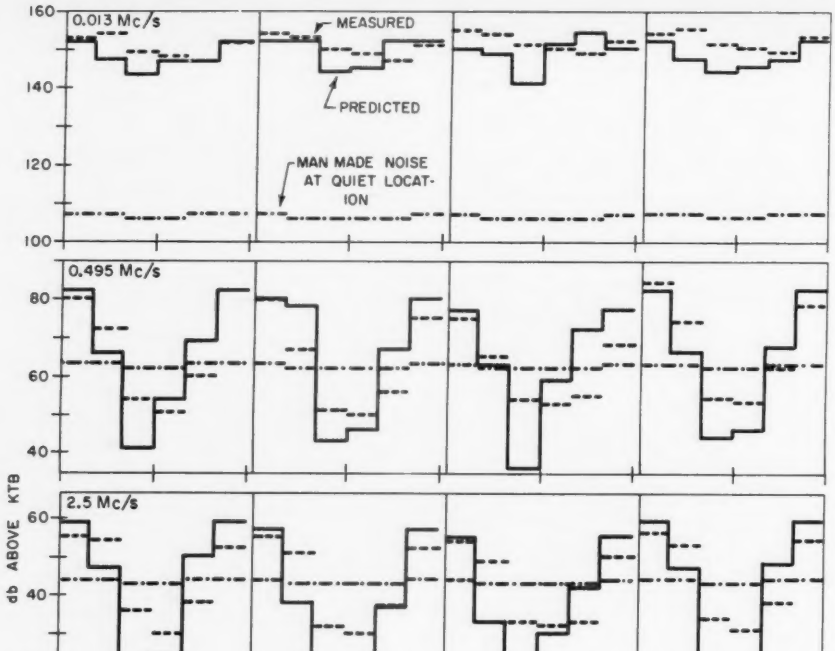

FIGURE 5. *Predicted and measured noise power at Kekaha, Hawaii.*
Prediction curves scaled from C.C.I.R. Report No. 65.

signals. In addition, as pointed out by Crichlow [1959 a], the methods used for determining antenna losses at Byrd have not been completely accurate, so this may contribute to the observed discrepancies between the predicted and measured noise values in figure 3. At times during winter a large amount of precipitation noise can be generated at the receiving antenna by snow being blown against the antenna.

Two additional noise sources may be of importance in polar regions, such as at Byrd Station. The first consists of aurorally generated noise which at times can be as much as 15 db above atmospheric noise on 18 Mc/s [Egan and Peterson, 1960] and significantly greater than atmospheric noise on 5 kc/s [Ellis, 1959]. The second, of importance on frequencies above about 2 Mc/s, is cosmic noise which can penetrate

the ionosphere near the poles where the critical frequency is low in winter, suffer ground reflection, and subsequently be propagated below the ionosphere by conventional means [Ellis, 1958].

To some extent the various types of noise sources which contribute to the measured values can be identified. For example, the presence of ionospherically propagated atmospheric noise in the Byrd Station data is indicated by the following two facts:

(1) A small seasonal variation can be seen (fig. 3) in the measured data on 0.051 and 0.545 Mc/s, where the minimum occurs during winter. This minimum probably corresponds to the absence of the *D* and *E* reflecting layers during the sunless Antarctic winter.

(2) The noise levels are generally higher in winter on the 5.0 and 10.0 Mc/s frequencies, complementing the behavior of the low frequencies in fact (1). That is, propagation of the high frequencies is expected to be better in winter because of the absence of the *D* region which is an absorber for these frequencies.

In considering the data for 0.051 and 0.545 Mc/s alone, one might be tempted to infer a predominance of manmade noise, where the seasonal minimum coincides with the relative inactivity of the station during winter. However, if this were true the noise level of the high frequencies would also be less during winter. On the other hand, that there is substantial manmade noise present is evidenced by an almost complete lack of seasonal and diurnal variation in the 2.5 Mc/s data. It should be noted that the measured values for this frequency fall about 20 db below the estimated level of manmade noise.

The measured noise values at Enkoping (fig. 4) on 0.545 and 2.5 Mc/s are substantially higher than predicted for all seasons during the daytime, which would indicate either a large contribution from manmade noise or interference, or an error of 20 to 30 db in the prediction. On 0.051 Mc/s the predicted values for the winter afternoon and evening time blocks are 21 and 18 db too high, respectively, and for the same time blocks during summer the predicted values are 16 and 5 db too low.

In figure 5 it can be seen that generally good agreement obtains between the predicted and measured noise values at Kekaha, Hawaii. During the day on 2.5 Mc/s the predicted noise level is on the order of 20 db below the measured level, but both are below the estimated manmade level and must be viewed with reservation. The higher-than-expected values measured on 5.0 Mc/s at Kekaha probably contain a large amount of station interference, as suggested by Crichlow [1959 a].

3.3. Month-Hour Median Measured Values

As an illustration of the diurnal noise patterns occurring seasonally in various parts of the world, month-hour medians of *Fa* measured on 5.0 Mc/s are plotted in figure 6 for selected months. The upper decile curves shown in this figure represent the noise value exceeded 10 percent of the time, while the lower decile value was exceeded 90 percent of the time.

Comparison of the month-hour median curves in figure 6 with the corresponding time-block curves in figures 2 to 5 reveal variations in the median of only 2 or 3 db within a 4-hr time block, except during sunrise and sunset periods. Month-hour median plots for the other frequencies compare similarly. With this in mind, the diurnal features of the different frequencies can be discussed on the basis of the curves in figures 2 to 5.

As expected, the diurnal variation of the frequencies at 0.495 Mc/s and higher show the effect of daytime ionospheric absorption, while the lower frequencies at 0.013 and 0.051 Mc/s exhibit little daily variation. The situation at Byrd Station, and to a lesser extent at Enkoping, is somewhat different. Smaller diurnal variations occur during summer and winter on the higher frequencies because the long periods of daylight during summer allows absorption 24 hr a day, which is completely absent during the sunless winter. The days of the transition seasons, fall and spring, have hours of daylight as well as darkness, so the diurnal variation at Byrd Station during these months is greater.

At the equator the times of sunrise and sunset do not change with season, so that the spread of the measured noise value in the 0400–0800 and 1600–2000-hr time blocks should be about the same all year. At Kekaha (22° N) the sun rises at about 0530 in summer and at about 0630 in winter, so this difference is not too important. However, at Enkoping (60° N) the sun rises at 0300 in summer (June) and not until 0900 in winter (December).

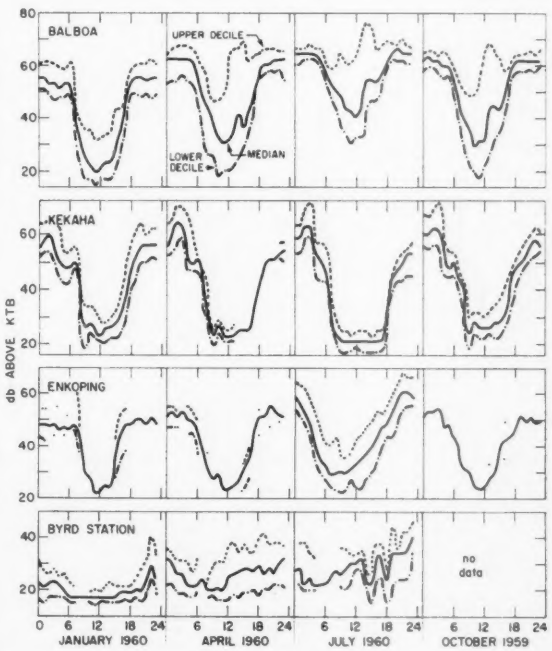

FIGURE 6. *Month-hour medians (measured) for selected months on 5.0 Mc/s.*

Thus, during summer at Enkoping large variations occur in the 0000–0400 hour time block (fig. 6), but during winter the variations due to sunrise conditions occur in the 0800–1200-hr time block. From this it is evident that the predicted values for time blocks containing sunrise and sunset can give only a rough indication of the median noise to be expected over the entire periods, and are less representative for individual hours than those for other time blocks.

3.4. Measured Variations in Medians

Within a 4-hr time block large day-to-day variations from the median are observed in the measured noise power. The noise power variations are expressed in terms of the ratio (in db) of upper decile to median, D_u, and lower decile to median, D_l, and are tabulated by Crichlow [1959 b] along with the monthly median values. An attempt was made in the C.C.I.R. report to predict the magnitude of D_u and D_l in each 4-hr time block as a function of frequency, but without regard to season or geographical location.

Thus, to compare the measured and predicted values of D_u and D_l the measured data were averaged over a 1-yr period (September 1959 to August 1960) in each 4-hr time block. For illustration, the average variations measured in the 1200–1600-hr time block at each of the 4 stations on the different frequencies are plotted in figure 7 as a function of frequency. The predicted values of D_u and D_l as a function of frequency for the same time block are shown as solid curves.

In figure 7 it can be seen that the measured variations are as much as 13 db greater than predicted (as at Balboa on 2.5 Mc/s) or 13 db less than predicted (as at Byrd Station on 0.545 Mc/s). At Kekaha, the predicted and measured values of D_u and D_l are generally in good agreement. Differences between the predicted and measured values of D_u and D_l for individual months are greater than 13 db in some cases. Generally, it can be said that the measured variations at high latitude stations are less than predicted while those at stations near the equator are greater than predicted.

3.5. Noise Power as Function of Frequency

Figure 8 shows the decrease in noise power with increasing frequency. The plotted points represent the measured noise power in the 0000–0400 time block averaged over 1 yr. The solid curve for each station was obtained from figure 22 in the C.C.I.R. report and indicates the expected variation of noise power with frequency.

The measured noise power variation with frequency as shown in figure 8 corresponds rather well with that predicted at all stations. The largest discrepancy occurs at Balboa on 0.013 Mc/s, where the measured value is about 17 db higher than that

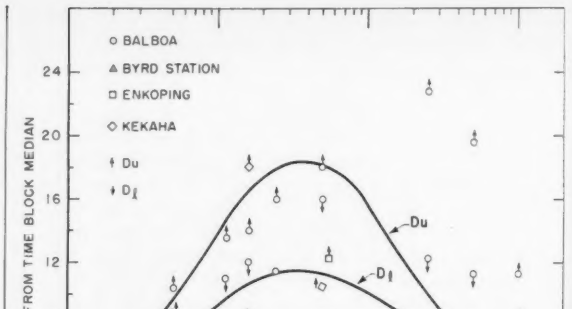

FIGURE 7. *Variations from time block median radio noise levels.*

Solid curves Du and D_l are upper and lower deciles predicted by C.C.I.R. Report No. 65, figure 25.

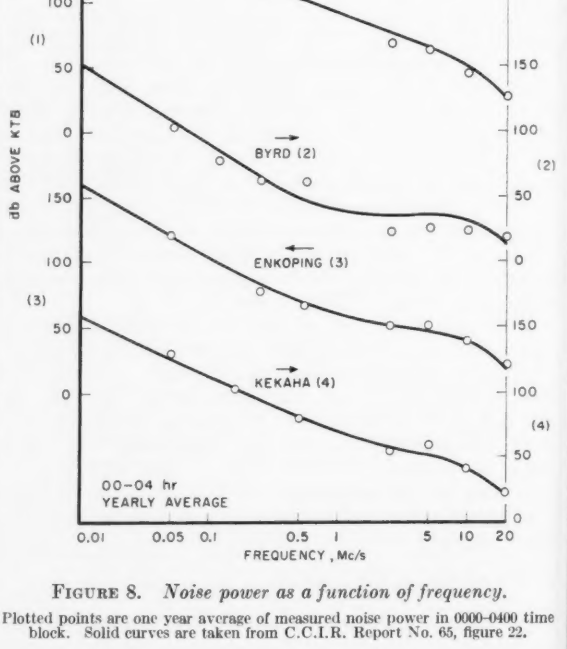

FIGURE 8. *Noise power as a function of frequency.*

Plotted points are one year average of measured noise power in 0000–0400 time block. Solid curves are taken from C.C.I.R. Report No. 65, figure 22.

predicted. Discrepancies greater than 10 db occur in only 10 percent of the measured points in figure 8. However, during individual months greater de-

partures appear, which should be taken into account in any future prediction curves. The good agreement between predicted and measured values in figure 8 indicates that the measured data considered here primarily represent noise of atmospheric origin.

3.6. Ratio of Measured to Predicted Values

The percentage distribution of the deviations of predicted from measured values are shown in figures 9a and 9b. The ratios of measured to predicted values are divided into the class intervals ± 0 to 10 db, ± 10 to 20 db, and ± 20 db or greater. The percentage distributions are shown in two ways. In figure 9a, the observed deviations for 1 yr counted on all frequencies at each station are given to illustrate the reliability of the predictions at different locations. In figure 9b, the observed deviations for the same year counted at all stations on each frequency are given to illustrate the reliability of the predictions for different frequencies.

From the percentage distributions of the ratios of observed to predicted noise levels given in figure 9a, it appears that the predictions are most reliable at Kekaha and Balboa, where over 80 percent of the observed deviations are 10 db or less. At Enkoping and Byrd Station less than 70 percent of the observed deviations are 10 db or less. At Byrd Station, 18.4 percent of the observed deviations were greater than ± 20 db, and nearly all of the observations were on the plus side. This strongly suggests local interference, which also adversely affects the percentage distribution of 0.545 Mc/s given in figure 9b.

The percentage distributions given in figure 9b indicate that the predictions are most reliable for 0.013 and 10.0 Mc/s, where 100 and 90 percent of the observed deviations were 10 db or less, respectively. The large percentage of deviations greater than 20 db (33%) observed on 0.545 Mc/s is due to the Byrd Station data. With those data excluded, the percentage deviation of the three class intervals for 0.545 Mc/s become 63, 21, and 16 percent, respectively.

The distributions for 2.5 and 5.0 Mc/s can be compared with a similar reliability test of the revised R.P.U. predictions as reported in Circular 462 (NBS, 1948). The comparison is not strict, however, due to the difference in sampling used here. The present distributions are based on the measurements of four stations, while those in Circular 462 were based on 17 stations. Thus, it is not unexpected that the reliability indicated in figure 9b on both frequencies is less than in Circular 462, where the percentage of observations for the three 10 db class intervals (0 to 30 db) were: 75, 20, and 5 percent for 2.5 Mc/s; and 90, 8, and 2 percent for 5.0 Mc/s.

4. Summary

Although good agreement is generally found between the predicted and measured noise power values, large differences still exist at some places and times. At places such as Kekaha, where the predictions are based on previous noise measurements, the differences are small, but at locations where the predictions are based on interpolated or extrapolated data (as at Byrd Station), large differences occur. The differences between predicted and measured values range from +40 to −24 db.

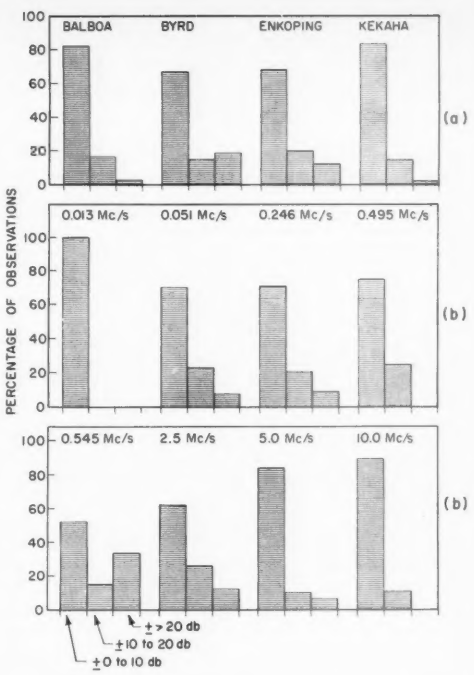

FIGURE 9. *Percentage of observed deviations: (a) on all frequencies at each station, and (b) at all stations on each frequency.*

In time blocks containing sunrise or sunset, the predicted values give only an indication of the median noise power to be expected over the entire periods and are less representative of individual hours than those for other time blocks.

The measured variations from the median, averaged over a 1 year period, differ from the predicted variations by as much as ± 13 db. In individual months larger discrepancies than this occur.

The noise power data considered in this paper follow the expected frequency dependence very well. Discrepancies greater than 10 db occur in only 10 percent of the cases, and the greatest departure is 17 db at Balboa on 0.013 Mc/s.

Of four stations investigated, it appears that the C.C.I.R. predictions are most reliable for Kekaha and Balboa, where over 80 percent of the ratio of measured to predicted values were 10 db or less, while less than 70 percent of the observations at Enkoping and Byrd Station were in this class. In comparing the predicted to measured values as a function of frequency, it was found that the ratio was 10 db or less in 100 and 90 percent of the observations for 0.013 and 10.0 Mc/s, respectively.

The author thanks R. Penndorf, S. C. Coroniti, and G. E. Hill for their critical reading of the manuscript and helpful suggestions, and Mrs. Carolyn Kotce for her tireless efforts in preparing the prediction curves for each station.

NOTE: Another interpretation of the discrepancies with the C.C.I.R. predictions will appear in a later issue of this journal under the authorship of W. Q. Crichlow and R. T. Disney of the Central Radio Propagation Laboratory, National Bureau of Standards. (Editor)

5. References

C.C.I.R. Report No. 65, Revision of atmospheric radio noise data, International Telecommunication Union, Geneva (1957).

Crichlow, W. Q., D. F. Smith, R. N. Morton, and W. R. Corliss, World-wide radio noise levels expected in the frequency band from 10 kilocycles to 100 megacycles, U.S. Dept. of Commerce, NBS Circ. 557 (Aug. 1955).

Crichlow, W. Q., C. A. Samson, R. T. Disney, and M. A. Jenkins, Radio noise data for the International Geophysical Year July 1, 1957—December 31, 1958, U.S. Dept. of Commerce, NBS Tech. Note 18 (July 27, 1959a).

Crichlow, W. Q., C. A. Samson, R. T. Disney, and M. A. Jenkins, Quarterly radio noise data, U.S. Dept. of Commerce, NBS Tech. Note 18 Series (1959b).

Egan, R. D., and A. M. Peterson, Auroral noise at HF. J. Geophys. Research **65,** 3830 (1960).

Ellis, G. R. A., The trapping of cosmic radio waves beneath the ionosphere, J. Atmospheric and Terrest. Phys. **13,** 61 (1958).

Ellis, G. R. A., Low frequency electromagnetic radiation associated with magnetic disturbances, Planetary and Space Science **1,** No. 4, 253 (1959).

National Bureau of Standards Circ. 462, Ionospheric radio propagation, U.S. Dept. of Commerce (1948).

Radio Propagation Unit Technical Report 5 (1st ed.), Minimum required field intensities for intelligible reception of radio telephony in presence of atmospherics or receiving set noise, Holabird Signal Depot, Baltimore, Md. (Dec. 1945); R.P.U. Technical Report 140 (revised ed., July 1947).

(Paper 65D6—160)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Effects of the Ionosphere on VLF Navigational Aids[1]

W. T. Blackband

Contribution from Radio Department, Royal Aircraft Establishment, Farnborough, Hants, England

(Received November 8, 1960; revised June 30, 1961)

Long range navigational aids could be devised using VLF transmissions. The propagation of such waves is controlled by the lowest level of the ionosphere. The small changes in phase velocity which accompany the diurnal ionospheric changes have been studied using ground monitor stations. Preliminary measurements made in an aircraft show internal consistency in fixing of about 1 nautical mile at ranges of 5,000 to 6,000 miles.

1. Introduction

The very low frequencies are those between 3 and 30 kc/s. This paper will not be concerned with any specific navigational aids but will discuss those factors of VLF propagation which will dominate the performance of any aids operating in this frequency band. It must be stressed that this description of the propagation should not be extrapolated into adjacent bands, and in particular would not apply without some considerable modification to the frequency band 70 to 120 kc/s.

In the VLF waveband the absorption at the ionosphere on reflection at large angles of incidence is small at all times of day and consequently propagation is possible even in daylight at 16 kc/s under circumstances in which there would be strong ionospheric absorption a few decades higher in frequency. For this reason propagation losses will not be discussed in detail and attention will be concentrated on the phase velocity of propagation of the signals because it is this velocity which determines the measured range from the transmitter in any phase comparison navigational aid. This velocity determines the wavelength which is the unit in which all measurements of distance are made. In general it is easier to observe changes in this velocity rather than determine its absolute value.

In discussion of VLF propagation we are concerned only with the lowest part of the ionosphere—the *D*-layer. The apparent height at nighttime is about 92 km while that at daytime is generally in the region of 72 to 80 km. For waves polarized with their electric vector normal to the ionosphere (vertically polarized), and at large angles of incidence, the signal is reflected without change in polarization, and the ionosphere acts as a sharply bounded conducting layer. Most VLF transmitting aerials radiate vertically polarized signals, and consequently there is little restriction in practice through confining attention in this paper to vertically polarized rays.

[1] This is a revised version of a paper presented at the Avionics Panel Meeting of the Advisory Group for Aeronautical Research and Development (AGARD), NATO, Istanbul, Turkey, October 1960.

2. Mechanisms of Propagation

The VLF band is of interest for navigational aids because of the long distance propagation which could permit the development of an aid of worldwide usefulness. The long-range propagation beyond the line of sight is achieved through repeated reflection of the transmitted wave between ionosphere and earth. This is illustrated in figure 1 which shows how, although the transmitter T is not directly visible from the receiver R, signals can be received having made one or two hops by way of the ionosphere which at these frequencies acts as a reflecting layer. Signals could reach R by a path involving many reflections—such as that shown by a dotted line, however because the reflection coefficient is small for small angles of incidence such signals would be rapidly attenuated, and the number of ray paths capable of transmitting an appreciable signal level to R is limited. It is usual to consider the propagation between earth and ionosphere in terms of the several ray paths for propagation up to 1,200 or 1,500 km. By taking into account the reflection coefficients of earth and ionosphere, it is possible by ray theory to explain [Wait and

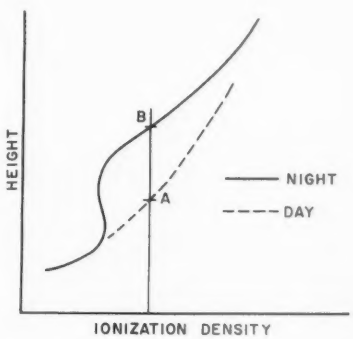

FIGURE 1. *Ionization contour for lower ionosphere (D region).*

Murphy, 1957] the main features of this medium-range propagation including the maxima and minima in signal corresponding to the Hollingsworth pattern of higher frequencies. These maxima and minima appear in figure 2 which shows a plot of field strength of GBR Rugby on 16 kc/s against great circle distance from the transmitter. This measurement was made in an aircraft flying at 40,000 ft from Farnborough, to Malta in daylight. This shows a minimum at about 230 nautical miles which arises through interference between the ground ray and first hop skywave of figure 1, and a second minimum at 740 nautical miles arising through interference between first and second hop skywaves.

At distances greater than about 1,200 km it is usual to describe propagation in terms of waveguide modes [Wait, 1957a; Budden, 1957]. The earth and ionosphere are then considered as the conducting walls of a waveguide. This waveguide differs from those in general use because these conducting walls are relatively far apart (about 2 ½λ at 10 kc/s and 8λ at 30 kc/s) and consequently several modes can be transmitted. If both earth and ionosphere were perfect conductors the distributions would be true sine waves. Thus we have the first mode which has a half sine wave distribution of voltage vector between ionosphere and earth, and the higher order modes which have distributions following several half sine waves. However, there are resistive losses in both earth and ionosphere and consequently the waves penetrate the waveguide walls. This results in a distortion from the true sine wave and some bending of the wavefront from the vertical. The losses in the ground and ionosphere are most marked for the higher order modes and consequently these modes are reduced to low amplitudes before they have traveled far. As a result of this the first mode usually dominates at distances of 2,000 km or more although for some ionospheric conditions the zero order mode may be the most important at the 3 kc/s end of the frequency band.

3. Attenuation of Signals

This is a facet of the subject of great complexity. Towards the lower end of the VLF band there is marked absorption while above 10 kc/s the attenuation is of the order of 2 db/1,000 km. In general the losses for transmission paths over sea are less than those for a corresponding overland path by a factor of as much as 2:1. A detailed theoretical study [Wait, 1957b] has been presented which expresses the frequency variation of attenuation for each mode in terms of the height of the ionosphere, and the conductivities of ground and ionosphere. Inasmuch as it is difficult to determine the properties of the lowest layers of the ionosphere except by VLF studies, it is not easy to confirm these theoretical predictions in detail. However, there is no doubt that they are a useful guide in the interpretation of the experimental phenomena. Figure 3 shows the measured variation of signal with distance for an over sea path from San Diego—Hawaii compared with that which would be predicted theoretically [Wait, 1957a] on the basis of chosen parameters for the ionosphere. The similarity between the forms of the two curves is pleasing, and the general agreement is as good as could be expected bearing in mind the somewhat arbitrary assumptions made about the ionosphere.

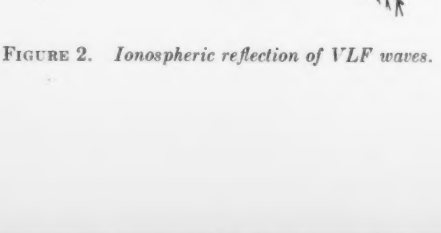

FIGURE 2. *Ionospheric reflection of VLF waves.*

4. Velocity

In any navigational system at present contemplated the observations will take the form of phase comparisons made either between a number of received signals, or between received signals and a high stability oscillator carried by the aircraft. For this reason we are interested in the phase velocity of the total wave. At distances from the transmitter for which the waveguide mode theory is applicable it is to be expected that the phase velocity would be somewhat greater than unity because of the finite spacing of ground and ionosphere, on the other hand, the effect of losses in the waveguide walls would be to slow down the wave. Some measured values have been published.

In the region near to the transmitter the effective phase velocity at any range can be deduced from the variation with radial distance of the resultant phase of the ground and sky rays, allowance being made for their differing ray paths and the changes of phase and amplitude on reflection.

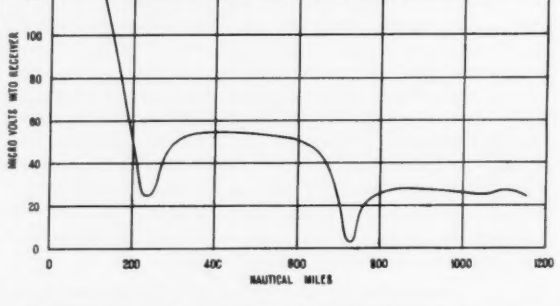

FIGURE 3. *Variation of signal amplitude with distance.*

5. Diurnal Changes in Phase Over a Given Path

If the phase of a signal received at a distant point is compared with that of a standard oscillator, then changes in velocity of propagation will become apparent as variations in relative phase. It has been found that there are diurnal variations in the phase of the received signal associated with changes from daylight to darkness. These are well illustrated in figure 4 in which J. A. Pierce [1957] depicts his "phase trapezium." The far left of the figure refers to times at which the whole 5,200-km path between the transmitter at Rugby and the receiver at Cambridge, Mass., is in darkness. At a short time before sunrise at Rugby a steady phase change commences, ending as the sun rises at Cambridge. During the day the phase remains constant until sunset at Rugby when the reverse process begins until the steady night value is reached at Cambridge sunset. The phase change is in a sense corresponding to a daytime increase in phase velocity. It is explained as the result of a reduction of the apparent height of the ionosphere during the day. This would be accounted for on the ray theory by saying that with the lower ionosphere the sky rays had a shorter path, or on the basis of waveguide modes, that the phase velocity had been increased by bringing the waveguide walls closer together.

It may be noted that the sharp onset of these phase changes occurs not at ground sunrise but a little earlier, at the time at which the sun viewed from a height of about 90 km above the receiver would be seen to rise above the ozonosphere.

The phase change of figure 4 has been found to be predictable from day to day within a few microseconds—an uncertainty in prediction which would introduce a positional error of about 1 mile in a navigational aid on a transatlantic flight. A series of measurements of the diurnal phase changes of the signals received from GBR have been made at Malta. These have shown variations which were basically repeatable from day to day but were more complex than those reported by Pierce. The changes for one day is shown in figure 5 from which it will be noted that instead of the neat trapezium there is a slight reverse drift before dawn followed by an overswing of phase before the steady day time value is reached. It has been suggested that these phase swings arise through the beating of signals reflected from two layers or that they represent the variations in phase of the sum of several modes, each mode being differently affected by the changing ionosphere. It would be expected that such overswing would be most apparent on a path which lay roughly parallel to the dawn or dusk line, because for this dawn or dusk would occur almost simultaneously along the whole path, all parts then being in the same intermediate state between day and night. For other paths where at any time parts are in different intermediate states, the overswings are averaged out. Even if the path follows the dawn of dusk line the effect of overswings would not be expected to appear

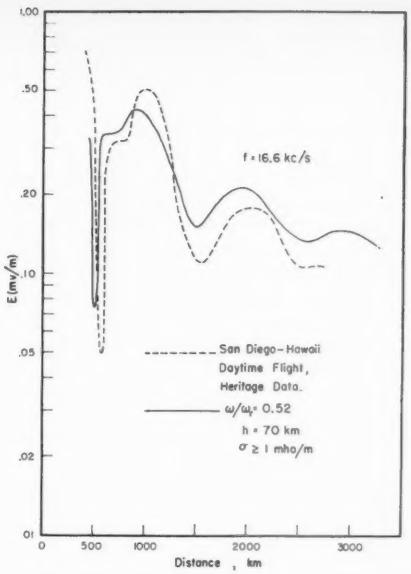

FIGURE 4. *Comparison of measured and computed field strengths (after Wail, Proc. IRE, June 1957).*

An experimental field-strength-versus-distance curve for 16.6 kc/s over sea and the corresponding calculated curve assuming $\omega/\omega_r=0.52$ or $\omega_r=2.0\text{X}10^5$, $h=70$ km and $\sigma\geq 1$ mho/meter.

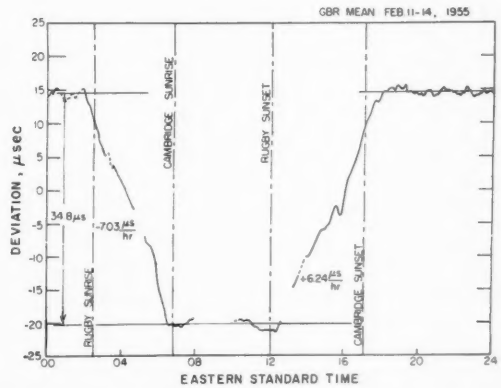

FIGURE 5. *The diurnal variation of the time of arrival of the GBR signal at Cambridge, Mass. (after Pierce, Proc. IRE, June 1957).*

at distances so great that only a single mode is received at an appreciable amplitude. At present it is not possible to check these theories because sufficient data are not available.

The magnitude of the diurnal phase change varies with distance from the transmitter. At distant receivers the ground ray will be negligible, and in order to compare signals near to the transmitter with those received at a distance it will be most convenient to neglect the ground ray and consider only the resultant sky ray.

At a point near to the transmitter the sky ray will be reflected at near normal incidence, and so when the reflecting layer changes by 18 km in apparent height from day to night the sky-ray path will change by 36 km—or a time of say 120 μsec. As the point of reception is moved further away from the transmitter the change in sky-ray path length will decrease rapidly at first and then more slowly as glancing incidence is approached. For greater distances the waveguide mode theory would predict a linear increase in phase change with distance. The initial rapid fall and linear rise would combine to give a variation of phase change with distance of the form of figure 6. Various measured values of phase change have been plotted and show a reasonable agreement with the expected shape of the curve. If this curve can be established by further measurement at other ranges up to say 10,000 km, then the diurnal variations of phase for a VLF navagational aid can be predicted with accuracy.

Such programs of monitoring diurnal changes have been planned during the summer of 1961. It is hoped to monitor simultaneously over 10 to 20 paths of a variety of distances. Some of these paths will lie entirely in the middle latitudes while others will be entirely tropical and others will cross the Arctic icecap.

6. Effects of Ionospheric Disturbances

VLF propagation can be affected by a magnetic storm or by a "sudden ionospheric disturbance," (SID). The effect of the magnetic storm is to give an increase in the fluctuations of phase during the night but no serious decrease in amplitude. Fluctuations of up to 5 μsec at 16 kc/s lasting 2 to 3 min have been reported [Pierce, 1957]. Following the main part of the storm the ionosphere remains disturbed and the diurnal effects are less than normal, recovering over a period (up to a few weeks in some cases). A recording of phase against time during a magnetic storm is shown in figure 7. This was over the path Rugby—Malta (2,210 km) for a frequency of 16 kc/s. It will be noted that the steady day phase is little different from that at night, but that the phase swings have the amplitude expected for a normal day. The features of this record are not yet explained although it is expected that the presence of the overswings at a time when the diurnal changes were small will be significant in the explanation of the overswings on normal days. It should be noted that although the effect of the magnetic storm in this case was to decrease the diurnal effect, there are circumstances in which the effect may be the reverse.

During the solar event of November 1960 the diurnal changes at Malta showed even less change from steady night to steady day, but the overswings were of full amplitude. A simultaneous phase record of GBR taken at Nairobi showed a trapezium of full amplitude. This diurnal change recorded at Nairobi is shown plotted as point 8 of figure 6. It will be noticed that it lies well on the curve, and is unaffected by the magnetic disturbance.

In order to assess the effects of magnetic storms on the accuracy of a VLF navigational aid, it will be necessary to obtain a wider range of test data by simultaneous and continuous recording on a world-wide basis. This study must determine how widespread and geographically uniform are the effects of a storm, and in particular, how the curve of figure 6 is altered. It may be that the effects of magnetic storms will best be countered by the use of an extra received signal which would help to establish and allow for changes in scale.

Compared with a magnetic storm an SID is of much less consequence. It is associated with a solar flare and effects only the sunlit hemisphere. As far as VLF propagation is concerned it alters the phase change on reflection at the ionosphere so as to give an advance in phase. The effect reaches its maximum in a few minutes and then decreases more slowly to regain a normal level within an hour. In general the amplitude of these changes for a given path is much less than the corresponding diurnal variations, but further data are required, and these could be expected as a byproduct of any monitoring such as that suggested for the study of magnetic effects.

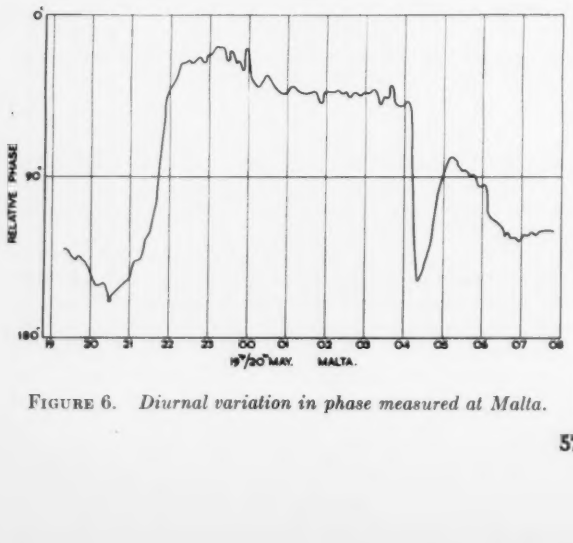

FIGURE 6. *Diurnal variation in phase measured at Malta.*

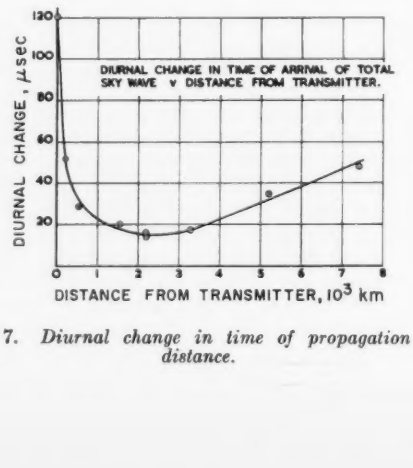

FIGURE 7. *Diurnal change in time of propagation versus distance.*

7. Airborne Measurements

In order to assess the potentialities of a VLF navigational aid a series of flights have been made by the Royal Aircraft Establishment, Farnborough. For this work a high stability oscillator was fitted in a Comet aircraft together with receivers which could compare the phase of this oscillator with the signals received from the transmitters GBR (Rugby, 16.0 kc/s) and GBZ (Criggion, 15.2 kc/s).

Suppose that the aircraft is directly above one of the transmitters and the phase of oscillator and transmitter compared, then if the aircraft moves away the phase of the transmitter will recede because of the time taken for the signal to pass to the aircraft. The recession of transmitter phase is a measure of great circle distance from the transmitter. If the aircraft is on a straight course at a steady speed, a plot of phase change against time will be a smooth curve. In figure 8 is shown a plot of such phase changes measured for GBR and GBZ. The slight scatter of the measured points about the smooth curve is equivalent to an uncertainty in great circle range of 1 nautical mile. Using these two transmitters this consistency has been obtained at distances up to 1,200 nautical miles from the transmitter.

Further airborne experiments have been carried out comparing the phases of the signals from GBR and NBA (Balboa, 18.0 kc/s). This has shown that it is possible to track the changes in position of the aircraft using changes in ranges from the two stations. Much of the flying has been over the Mediterranean, and this has shown that the fixes obtained at ranges between 5,000 and 6,000 miles from Balboa on any given flight are internally consistent to within a nautical mile. Their absolute accuracy is not easy to determine because of the difficulty of locating the position of a high flying aircraft. However, a further series of flights is planned during which time theodolite and other photographic methods of determining position will be used.

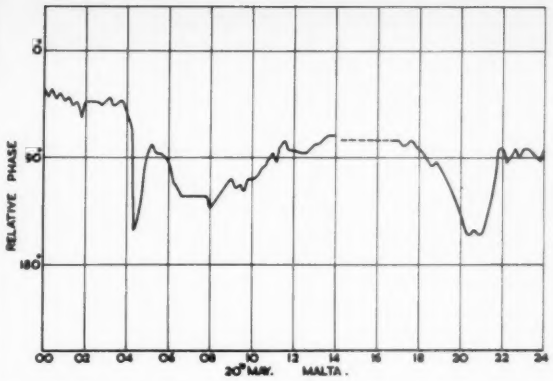

FIGURE 8. *Diurnal variation of relative phase during a magnetic storm.*

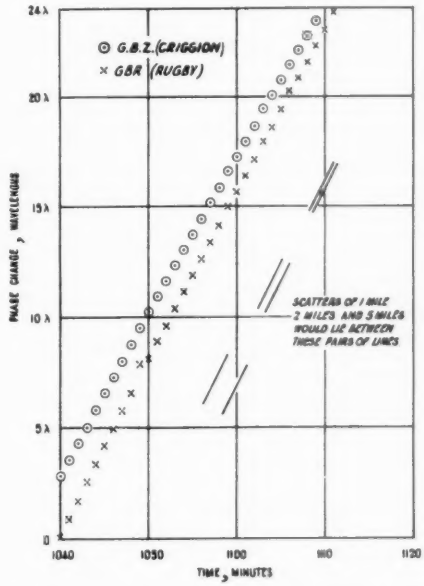

FIGURE 9. *Variation in measured phase against time for a flight on a straight course.*

8. Conclusion

The general picture of ionospheric effects on VLF propagation is that there are no fundamental reasons why worldwide navigational aids could not be devised in this frequency band. However, there is a lack of data to fill in the detail of the picture, and only on the basis of such details could one assess the ease and convenience of navigation using such a system. In this field, theory has outstripped experiment, and the need is for an extensive and worldwide monitoring of the phase stability of VLF transmissions over long paths during both normal and disturbed ionospheric conditions.

9. References

Budden, K. G., The waveguide mode theory of the propagation of very-low-frequency radio waves, Proc. IRE **45,** No. 6, 772–774 (June 1957).

Pierce, J. A., Intercontinental frequency comparison by very-low-frequency radio transmission, Proc. IRE **45,** No. 6, 794–806 (June 1957).

Wait, J. R., The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. IRE **45,** No. 6, 760–767 (June 1957a).

Wait, J. R., The attenuation vs. frequency characteristics of VLF radio waves, Proc. IRE **45,** No. 6, 768–771 (June 1957b).

Wait, J. R., and A. Murphy, The geometrical optics of VLF sky wave propagation, Proc. IRE **45,** No. 6, 754–760 (June 1957).

Additional References

Bain, W. C., R. N. Bracewell, T. W. Straker, and C. H. Westcott, The ionospheric propagation of radio waves of frequency 16 kc/s over distances of about 540 km, Proc. Inst. Elec. Engrs. **99,** pt IV, No. 3, 250–459 (July 1952).

Bracewell, R. N., The ionospheric propagation of radio waves of frequency 16 kc/s over distances of about 200 km, Proc. Inst. Elec. Engrs. **99,** pt IV, No. 3, 217–228 (July 1952).

Bracewell, R., K. G. Budden, J. A. Ratcliffe, T. W. Straker, and K. Weekes, The ionospheric propagation of low- and very-low-frequency radio waves over distances less than 1,000 km, Proc. IRE **98,** pt III, 221 (1951). This is an integrating paper surveying the work of the teams working in conjunction with the Cavendish Laboratory at Cambridge.

Wait, J. R., A survey and bibliography of recent research in the propagation of VLF radio waves, NBS Tech. Note 58 (May 1960).

Waynick, A. H., The present state of knowledge of the lower ionosphere, Proc. IRE **45,** No. 6, 741–749 (June 1957).

Yabroff, I. W., Reflection at a sharply bounded ionosphere, Proc. IRE **45,** No. 6, 750–753 (June 1957).

(Paper 65D6–161)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

On the Spectrum of Terrestrial Radio Noise at Extremely Low Frequencies[1,2]

H. R. Raemer

Contribution from Applied Research Laboratory, Sylvania Electronic Systems, Waltham, Mass.

(Received May 30, 1961; revised June 15, 1961)

A theory of the frequency spectrum of radio noise at extremely low frequency (ELF) is presented and the results compared with recent measurements of the first five "Schumann" resonant modes (between 8 and 34 c/s) made by Balser and Wagner [1960]. The source of this noise is assumed to be return strokes in vertical cloud-ground lightning flashes distributed randomly in time, uniformly in angular displacement along the earth relative to the observer and with statistics of stroke duration, interstroke intervals and strokes per flash taken from studies of thunderstorms reported by J. C. Williams. Thus, the mathematical model for the noise sources is an extremely simple one, being analogous to the shot effect in electron devices. The electromagnetic model employs the familiar waveguide mode theory, assumes a sharply bounded homogeneous ionosphere, and neglects the earth's magnetic field.

Agreement between the shape of the theoretical and observed spectrum is good for the first three modes and rather poor for the higher modes.

It is found by matching the theoretical resonant frequencies to the observed resonances that the product of effective ionosphere height h and the square root of effective conductivity $\sqrt{\sigma_i}$ is a decreasing function of frequency. The functional dependence of this quantity on frequency is determined and used in the calculation of the mode spectrum.

Discrepancies between the theory and experimental results are believed to be partially due to the artificiality of the sharply bounded homogeneous ionosphere model and to failure to give sufficient probability weighting to equatorial regions of abnormally high thunderstorm activity.

These last items are the subjects of continuing work on the extension of the theory.

1. Introduction

The theory of extremely low frequency (ELF) propagation, as originally developed by Schumann [1957] and extended by Wait [1960a] and others, predicts the existence of a series of modes whose resonant frequencies, relative amplitudes, and Q's can be found by considering the earth-air-ionosphere cavity as a linear filter and calculating its frequency response function directly from electromagnetic theory. Recently, Balser and Wagner [1960] have observed these modes and experimentally determined the detailed shape of the mode spectrum.

The basis of the cavity analog is the classical Sommerfeld theory of propagation around a spherical earth, which leads to the representation of the field between earth and ionosphere as a series of "waveguide modes." If the ionosphere height is extremely small compared to wavelength, as is true at ELF, the electric field is essentially radial and a single waveguide mode is sufficient to describe the fields at a distance from a localized source. This waveguide mode when considered in the time domain can, in turn, be shown (by an appropriate expansion of the hypergeometric series into a series of Legendre polynomials) to be a superposition of waveforms each closely approximating a damped sinusoid of frequency $\frac{1}{2\pi}\frac{c}{a}\sqrt{n(n+1)}$, where c and a are velocity of light and earth radius, respectively and n is a mode index number running from zero to infinity. This phenomenon is entirely analogous to the ringing of a resonant cavity in response to impulse excitation.

Experimentally observed VLF noise is thought to consist of a superposition of responses from lightning flashes all over the world [Williams, 1959]. Therefore, assuming that this hypothesis is correct down to ELF, a theoretical model based on it should yield accurate

[1] Work supported by the Office of Naval Research.
[2] A preliminary note on this work appeared in the J. Geophys. Research, May 1961.

calculations of the spectrum of observed ELF noise. It is the purpose of the study reported here to construct such a model and use it to calculate the expected ELF noise spectrum. The results are compared with the measurements of Balser and Wagner [1960].

2. ELF Cavity Mode Theory

Following Wait [1960a] one can express the significant (radial) component of the electric field at an observation point due to the harmonic excitation (at frequency ω) of a vertical electric dipole of moment Ids on the surface of the earth in the form

$$E_r(i\omega)=Ids\sum_{n=0}^{\infty}C_n(\theta)F_n(i\omega) \tag{1}$$

where the ionosphere is assumed to be electrically uniform and sharply bounded at height h, and the following definitions apply:

$$C_n(\theta)=\frac{i}{4\pi a^2\epsilon_0 h}P_n(\cos\theta)(2n+1)$$

ϵ_0=Dielectric constant of free space

$P_n(\cos\theta)$=Legendre polynomial of order n

θ=Polar angle relative to the source

and

$$F_n(i\omega)=\frac{1}{i\omega}\frac{-\omega_v^2(i\omega)}{[-\omega_v^2(i\omega)+\omega_n^2]}$$

with

$$\omega_v^2(i\omega)=\frac{c^2}{a^2}v(v+1)=-(i\omega)^2-\alpha(i\omega)^{3/2}$$

$$\omega_n^2=\frac{c^2}{a^2}n(n+1)$$

and

$$\alpha=\frac{1}{h\sqrt{\mu\bar{\sigma}_i}}\left(1+\sqrt{\frac{\bar{\sigma}_i}{\bar{\sigma}_g}}\right)$$

where

$$\bar{\sigma}_g=\sigma_g+i\epsilon_g\omega_g\approx\sigma_g=\text{complex conductivity of the earth}$$

and

$$\bar{\sigma}_i=\frac{e^2}{m}\left\{\frac{N}{(g-i\omega)}\right\}=\sigma_i+i\omega\epsilon_i\approx\sigma_i=\text{complex conductivity}$$

of the ionosphere, where N, e, m, g, are the electron density, charge, mass, and collision frequency [Stratton, 1941], and it is assumed that $\omega << g$. In the calculation of the appropriate ionospheric electrical parameters the effects of the earth's magnetic field and the motions of the heavy ions are neglected; in addition, the electron collision frequency is taken to be independent of the nonrandom electron energy. In general $|\bar{\sigma}_g|>>|\bar{\sigma}_i|$, so in what follows α, which will be called the "ionospheric loss parameter," is approximately a real quantity, equal to $1/(h\sqrt{\mu\sigma_i})$.

The function $E_r(i\omega)$ is actually the Fourier transform of the field in response to a (vertical) current impulse of magnitude I at $\theta=0$ and at $t=0$, i.e., it can be regarded as the characteristic frequency response function of the earth-air-ionosphere cavity. $F_n(i\omega)$ is the Fourier transform of the nth mode excited by an impulse of dipole current of magnitude I. From knowledge of all of the $F_n(i\omega)$'s one can predict the frequency response to any time varying source current using standard linear system theory.

3. Superposition of Fields From ELF Sources

The radial field component is expressed here in a spherical coordinate system wherein the source is at the pole. Because of azimuthal symmetry about the source, we can without changing the form of (1) consider $E_r(\omega)$ as expressed in a system with pole at the observer and source at polar angle θ. Suppose that lightning flashes, assumed equivalent to vertical dipole sources, are distributed around the earth's surface. We define $E_r(\omega;\theta,\phi)$ as the Fourier transform of the radial field component at the pole due to a vertical dipole source at the point θ,ϕ. The source current moment waveform at θ,ϕ and its Fourier transform are denoted by $\dot{i}(t;\theta,\phi)$ and $\underset{\sim}{I}(\omega;\theta,\phi)$, respectively.

The voltage induced in a receiver at the observation point at time t due to the presence of lightning sources all over the world is proportional to

$$v(t)=\sum_{k=1}^{N}\int_{-\infty}^{t} dt' f(t-t';\theta_k,\phi_k)\,\dot{i}(t';\theta_k,\phi_k) \tag{2}$$

where the index k is used to number the source positions, from 1 through N, and

$$f(t;\theta,\phi)=\sum_{n=0}^{\infty}(2n+1)P_n(\cos\theta)f_n(t)$$

where $f_n(t)$ is the inverse Fourier transform of $F_n(i\omega)$, i.e., the impulse response of the *nth* earth-air-ionosphere cavity node.

4. Calculation of the Power Spectrum

The evaluation of the power spectrum of $v(t)$ falls within a familiar class of problems in random noise theory, specifically, the spectral analysis of the voltage arising from the "shot effect" in electron devices. This problem has been treated by Rice [Wax, 1954] and others [Davenport and Root, 1958]. The analysis to follow is mathematically equivalent to that of the shot effect.

Consider a voltage, $v(t)$, appearing at the output of a linear filter with impulse response $f(t)$ due to a train of input pulses, each of which is denoted by $\xi(t-t_k;\beta)$. The shape of the *kth* pulse is determined by the components of a vector β, its magnitude (time integral of amplitude over duration of pulse) is a_k, and its time of occurence is t_k. The output voltage is thus

$$v(t)=\sum_{k} a_k\int_{-\infty}^{\infty} f(t-t')\xi(t'-t_k;\underset{\sim}{\beta_k})dt' \tag{3}$$

We assume a time interval [3] $-T\leq t\leq T$ over which all the action takes place and that there are exactly N pulses within that interval. Because of finite filter "memory" and the filter realizability requirement,[4] we can regard as zero the response to pulses occurring at values t_k outside the interval $(t-T_0)$ to t. Here T_0 is the time of occurence of the earliest imput influencing the filter output at time t. The interval $(t-T_0)$ is less than T but somewhat greater than the filter time constant. Numbering from 1 to N the pulses within the interval $-T$ to T, and using (3) we can write

$$v(t)=\sum_{k=1}^{N} a_k g(t-t_k;\underset{\sim}{\beta_k}) \tag{4}$$

where

$$g(t-t_k;\beta_k)=\int_{t-T_0}^{t} f(t-t')\xi(t'-t_k;\underset{\sim}{\beta_k})dt';\ -T\leq t_k\leq T;\qquad k=1,\ldots,N,$$

the response of the filter at time t to a pulse occurring at time t_k.

---

[3] The time T washes out in the final results; the finite time assumption is strictly for mathematical convenience in the derivation.

[4] i.e., the requirement that the filter can respond only to past (and not to future) inputs.

We now assume that: (1) The a_k's are independent and identically distributed random variables with common probability density function $p(a)$; (2) for a fixed number of pulses within the interval $t-T_0$ to t, the t_k's are statistically independent random variables, each uniformly distributed throughout the interval $-T$ to T; (3) the probability of finding exactly N pulses within the interval $t-T_0$ to t, denoted by $p(N)$, is independent of the statistics of the a_k's and β_k's; (4) the β_k's are independent and identically distributed random variables with common probability density function $p(\beta)$; and (5) the random variables a_k, t_k, and β_k are all mutually statistically independent.

By calculating the autocorrelation function of $v(t)$ under the assumptions (1) through (5) and taking its Fourier transform, we obtain for the spectrum [5]

$$S(\omega)=\frac{1}{4T^2}\{\langle N(N-1)\rangle\langle a\rangle^2\overline{\langle g(t)\rangle\langle g(t+\tau)\rangle}\delta(\omega)\}+\hat{\bar{S}}(\omega) \tag{5}$$

where

$$\hat{\bar{S}}(\omega)=\frac{\langle N\rangle\langle a^2\rangle}{2T}\int_{\text{All }\underline{\beta}\text{-space}} d\underline{\beta}\,p(\underline{\beta})|G(\omega;\underline{\beta})|^2$$

and

$$G(\omega;\underline{\beta})=\int_{-\infty}^{\infty} dt e^{-i\omega t}g(t;\underline{\beta})$$

By the well-known convolution theorem for Fourier transforms

$$G(\omega;\underline{\beta})=F(i\omega)\tilde{\xi}(\omega;\underline{\beta}) \tag{6}$$

where $F(i\omega)$ is the frequency response function of the filter and $\tilde{\xi}(\omega;\underline{\beta})$ is the Fourier transform of the pulse.

The first term of (5) is a "DC spike," and is not required to determine the shape of the mode spectrum in the frequency region of interest. We will, therefore, concern ourselves only with $\hat{\bar{S}}(\omega)$.

5. Statistics of Lightning Source Currents

The result (5) can be adapted to the calculation of the ELF noise spectrum by allowing the pulses $\xi(t-t_k;\underline{\beta}_k)$ to represent lightning flashes whose shape parameters, time of occurrence t_k, and polar angles θ_k are assumed to be random variables, statistically independent when $j\neq k$. The angles θ_k are included among the components of the vector $\underline{\beta}_k$, as are the number of strokes in a flash [Williams, chapter VI], and the shaping parameters associated with individual strokes.

The statistics of the lightning flashes are assumed to obey all of the assumptions (1) through (5). It does not necessarily follow that we can with physical justification define a spatial probability density function $p(\theta)$ apart from consideration of other random variables associated with flashes. However, this assumption is basic to what follows and we note that it is tantamount to the assumption that statistics of lightning flash shape (i.e., those types of flash important as sources of VLF noise) are independent of those of terrestrial location of the flashes. The essential correctness of the latter assertion is supported by experimental evidence [Williams, tables 6–1, 6–2, 6–4, figures 6–1, 6–2, 6–7, 6–11, 6–14].

Using (5) and (6), the definition of $f(t)$ given in (2), and the assumptions about lightning flashes discussed above, we obtain for the ELF noise spectrum (exclusive of the DC component)

$$\hat{\bar{S}}(\omega)=K\langle|\tilde{\xi}(\omega,\underline{\beta}')|\rangle\sum_{m,n=0}^{\infty} S_{mn}(\omega) \tag{7}$$

where K is a constant, $\underline{\beta}'$ represents all lightning flash parameters exclusive of θ, and

$$S_{mn}(\omega)=(2m+1)(2n+1)F_m(i\omega)F_n^*(i\omega)L_{mn},$$

[5] Ensemble averaging over β_k is denoted by $\langle\ldots\rangle$, averaging over the time variables t_k by a line over the symbol.

where

$$L_{mn}=2\pi\int_0^{\pi} d\theta \sin\theta P_m(\cos\theta)P_n(\cos\theta)p(\theta),$$

$$\langle|\tilde{\xi}(\omega;\underline{\beta}')|^2\rangle=\int_{\text{All }\underline{\beta}'\text{ space}} d\underline{\beta}'p(\underline{\beta}')|\tilde{\xi}(\omega;\underline{\beta}')|^2$$

$p(\underline{\beta}')$ is the probability density of the set of parameters $\underline{\beta}'$, and $p(\theta)\sin\theta d\theta$ is the probability that the lightning flash is located between θ and $\theta+d\theta$.

We are interested only in relative and not absolute amplitudes; therefore, in numerical computations we will fix K by assuming a normalization that equates the average power levels in the theoretical and experimental results.

6. Flash Shape Parameters

J. C. Williams [ch. V, sec. 5f] has concluded from his studies of thunderstorms that the overwhelming cause of VLF noise is the presence of vertical currents in the return strokes associated with cloud-to-ground lightning flashes. These flashes consist of from 2 to 12 strokes of similar shape. He has compiled statistics on flash parameters, including the shape of the median return stroke [ch. VI, sec. 1, 4, 5, and ch. VII, sec. 1], interstroke interval, and number of strokes per flash. He has also made spectral calculations for the VLF noise resulting from the flashes, but these are in a higher frequency range and inapplicable to our problem (ch. VII, sec. 3).

The current moment of Williams' median return stroke is proportional to

$$i(t-t_k)=\left(\int_0^t \hat{v}(t')dt'\right) i(t-t_k)=\{-16.8(10^3(e^{-(t-t_k)/(1.7\times10^{-6})}+15.35(10^{-3})e^{-(t-t_k)/(33\times10^{-6})}$$

$$+10^3e^{-(t-t_k)/(5\times10^{-4})}+0.45(10^3)e^{-(t-t_k)/(6.8\times10^{-3})}\}u(t-t_k)\{1-e^{-5.55(10^{-3})(t-t_k)}\} \quad (8)$$

where $\hat{v}(t)$ is the velocity of the lightning charge along the cloud-ground channel, the unit step is inserted to assure that the stroke has zero value before time t_k and the lightning current $i(t-t_k)$, $v(t)$ and t_k are measured in amperes, kilometers/second and seconds, respectively.

The normalized spectrum of (8) is proportional to

$$|I(\omega)|^2=\left|e^{i\omega t_k}\sum_{l=1}^{4} I_l\tau_l\left\{\frac{1}{1-i\omega\tau_l}-\frac{\tau_v}{\tau_l+\tau_v-}\right\}\right| \quad (9)$$

where I_l and τ_l are respectively the currents and time constants given in (8) and $\tau_k=1.80(10^{-4})$ sec.

It would be desirable in the interest of accuracy to consider the shaping parameters of the stroke model as random variables and average the spectrum over these variables. This is a formidable mathematical problem in general; and moreover, there is insufficient experimental data available to assign probability density functions to these parameters, a measure which would be required to evaluate the averages in the general case. The assumption that all shaping parameters have small standard deviations relative to their mean values would reduce the problem to one requiring only a knowledge of first and second moments of the distribution of each parameter. However, experimental knowledge of even these quantities is not available in enough detail to allow any meaningful assignment of numbers to the calculation of the spectrum of the median return stroke itself. This is tantamount to assuming that all return strokes are nearly identical in shape. This is not quite accurate, but the analysis will still provide useful insights into the effects of stroke shape on the ELF spectrum.

The waveform of a flash consisting of M identical strokes [6] is

$$\xi(t)=\sum_{l=1}^{M} \underset{\sim}{i}(t-t_l), \tag{10}$$

where $\underset{\sim}{i}(t-t_l)$ is given by (8).

The spectrum of a lightning flash is roughly proportional to the averaged absolute square of the Fourier transform of $\xi(t)$, i.e.,

$$\langle|\tilde{\xi}(\omega)^2\rangle=|\underset{\sim}{I}(\omega)|^2 \sum_{l,k=1}^{M} \langle e^{i\omega(t_l-t_k)}\rangle \tag{11}$$

where the indicated average is over the ensemble of the interstroke time intervals t_l-t_k.

An obvious classification of terms in the summation of (11) results in

$$\rho_M(\omega)\equiv\frac{\langle|\tilde{\xi}(\omega)|^2\rangle}{|I(\omega)|^2}=M+\sum_{\substack{l,k=1\\ l\neq k}}^{M} e^{i\omega(l-k)T}\langle e^{i\omega\Delta T_{lk}}\rangle \tag{12}$$

where T is the average interstroke interval and ΔT_{lk} is the deviation of the interval (t_l-t_k) from $(l-k)T$.

To calculate $\langle e^{i\omega\Delta T_{lk}}\rangle$ accurately, we would require a knowledge of the probability density function for ΔT_{lk}. An extremely rough estimate of this function is provided by Williams [ch. VI, figs. 6–7], but it is not specified in sufficient detail to allow a precise representation. A rough processing of Williams' curve shows a Gaussian distribution to be a reasonable approximation for this variable, with standard deviation σ_{lk}, the latter being assumed somewhat smaller than T. The result of a simple integration is

$$\langle e^{i\omega\Delta T_{lk}}\rangle=e^{-\omega^2\sigma^2_{lk}/2} \tag{13}$$

In the absence of experimental data on second order statistics of ΔT_{lk}, it will be assumed that the random variable $(t_l-\langle t_l\rangle)$ is not strongly correlated with $(t_k-\langle t_k\rangle)$, where $k\neq l$. It follows that

$$\sigma^2_{lk}\simeq\sigma^2_l \qquad \text{for all } l\neq k \tag{14}$$

Invoking (13) and (14) in (12), using the power series for $1/(1-x)$ to bring the summation into closed form and averaging over all possible values of M, we arrive at the results.

$$\rho(\omega)=\sum_{M=0}^{\infty} \rho_M(\omega)\,p(M) \tag{15}$$

where $p(M)$ is the probability of M strokes in a flash and

$$\rho_M(\omega)=M(1-e^{-\omega^2\sigma_l^2/2})+e^{-\omega^2\sigma_l^2/2}\left(\frac{\sin^2\dfrac{M\omega T}{2}}{\sin^2\dfrac{\omega T}{2}}\right) \tag{16}$$

Averaging (16) over values of M, we obtain

$$\rho(\omega)\simeq\langle M\rangle(1-e^{-\omega^2\sigma_l^2/2})+e^{-\omega^2\sigma_l^2/2}\sum_{n=0}^{\infty} A_n \operatorname{cosn}\,\omega T \tag{17}$$

where

$$A_k=\sum_{M=K}^{\infty}(M-K)p(M).$$

[6] Actually the strokes are not quite identical, the first being about twice as intense as succeeding ones. However, the approximation will not introduce serious errors into an analysis of this type.

The parameters T and σ_t have been roughly estimated from Williams (fig. 6–7) and found to be about 33 and 7 msec, respectively. The precision of these estimates is questionable, but they are probably correct to within 20 percent. The same remarks apply to similar rough estimates made of $p(M)$ from Williams' cumulative probability distribution curves for the number of strokes per flash (fig. 6–1). With these estimates we conclude that the probabilities $p(1)$ through $p(10)$ can be roughly approximated by the numbers 0.15, 0.22, 0.23, 0.15, 0.09, 0.06, 0.04, 0.03, 0.02, 0.01, and 0, that $\langle M \rangle$ is about 2.5 msec, that the numbers A_0 through A_9 are roughly equal to 3.5, 2.5, 1.6, 1.0, 0.61, 0.36, 0.20, 0.10, 0.04, and 0.01, and that A_n effectively vanishes for $n>9$.

7. Geographical Distribution of Lightning Activity

The assumption is now introduced that the probability of occurrence of a lightning flash within a range of angles θ, $\theta+d\theta$ during the time interval between $t-T_0$ and t is

$$\sin\theta p(\theta)d\theta=\int_0^{2\pi} d\phi(p(\theta,\phi)\sin\theta d\theta)=(p_0+\bar{p}(\theta))d\theta\sin\theta; \qquad 0\leq\theta\leq\pi \tag{18}$$

where $p(\theta,\phi)d\theta\sin\theta d\phi$ is the flash occurrence probability for a region of solid angle $d\theta\sin\theta d\phi$ at θ,ϕ, p_0 is the part of the θ distribution that is uniform in the angle θ, and $\bar{p}(\theta)$ accounts for the existence of regions of abnormally high or low average thunderstorm activity.

Initially, we will neglect $\bar{p}(\theta)$ and assume a completely uniform distribution p_0. The result of applying this assumption to (7) (see appendix) with the aid of the orthonormality of the Legendre polynomials is

$$\frac{\hat{S}(\omega)}{p_0\langle|\bar{\xi}(\omega,\underline{\beta}')|^2\rangle}=\sum_{n=0}^{\infty}(2n+1)|F_n(i\omega)|^2 \tag{19}$$

Considering now the term $\bar{p}(\theta)$, it has been established that there are three equatorial regions of abnormally high average thunderstorm frequency [Handbook of Geophysics, 1960]. These regions are located in (1) South America, (2) Africa, and (3) the Southwest Pacific area. We will denote them respectively by R_1, R_2, and R_3. The regions center approximately at (50° W, −10° S), (15° E, +10° N), and (110° E, +10° N) respectively, and during all seasons of the year they show an average thunderstorm frequency somewhat higher than is observed elsewhere in the world. Speaking more quantitatively, the spatial probability density in those regions is from 100 to 1,000 times higher than that throughout most of the earth's surface.

There is another near-equatorial region in Southeast Asia (~90° E, +20° N), showing substantial year-round thunderstorm activity, but not as high as in R_1, R_2, and R_3. This region has not been accounted for in our calculations. Certain nonequatorial land regions show intense activity during their summers, e.g., the Southeastern United States (centered about (75° W, +30° N)), but this has also been excluded from our analysis. It is desired here to concentrate on those features of the ELF noise spectrum that are not season-dependent, and this objective is best fulfilled by studying the contributions from the three most prominent equatorial regions R_1, R_2, and R_3.

In view of the approximate nature of meteorological data, it would not be feasible to attempt to construct a precise thunderstorm probability function for the regions under investigation. We will, therefore, construct an idealized theoretical model which is convenient for computation and which embodies the essential features of the observed distributions.

We will approximate the probability density functions within the regions R_j by Gaussian functions

$$p_j(\theta)=K_j\frac{e^{-(\theta-\theta_j)^2/2\sigma_j^2}}{\sqrt{2\pi\sigma_j^2}} \tag{20}$$

where K_j is a constant, θ_j is the center of the distribution in R_j and σ_j is the standard deviation.

Using (7), (18), and (20), we have

$$L_{mn}=\frac{4\pi p_0}{2n+1}\,\delta_{mn}+2\pi\sum_{j=1}^{N}K_j l_{mnj} \tag{21}$$

where

$$l_{mnj}=\int_0^{\pi}d\theta\,\sin\theta P_m(\cos\theta)P_n(\cos\theta)\,\frac{e^{-(\theta-\theta_j)^2/2\sigma_j^2}}{\sqrt{2\pi\sigma_j^2}}$$

The integrals l_{mnj} have no simple orthogonality property to aid in their evaluation. The method to be used here will involve the expansion of the Legendre polynomials $P_n(\cos\theta)$ in a Fourier Series of the form [Jahnke-Emde, 1945]

$$P_n(\cos\theta)=\sum_{\mu=0}^{n}\tilde{a}_{n\mu}\cos\mu\theta \tag{22}$$

where the coefficients $\tilde{a}_{n\mu}$ are given in reference 9. From (21) and (22) we have

$$l_{mnj}=\sum_{\mu=0}^{m}\sum_{u=0}^{n}\tilde{a}_{m\mu}\tilde{a}_{nv}\mathsf{U}_{m\mu nv} \tag{23}$$

where

$$\mathsf{U}_{m\mu nv}\equiv\int_0^{\pi}d\theta\,\sin\theta\,\frac{e^{-(\theta-\theta_j)^2/2\sigma_j^2}}{\sqrt{2\pi\sigma_j^2}}\cos\mu\theta\cos\nu\theta$$

Using trigonometric identities, assuming (justifiably, with the parameter values of regions R_1, R_2, R_3) that (1) σ_j is small compared to π and (2) θ_j is not near 0 or π (so that the limits of the integral are effectively $-\infty$ and $+\infty$), and integrating by completion of the square in the exponential, we obtain

$$\frac{\hat{S}(\omega)}{p_0\langle|\tilde{\xi}(\omega,\underline{\beta}')|^2\rangle}-\sum_{n=0}^{\infty}(2n+1)|F_n(i\omega)|^2=\frac{1}{P_0}\sum_{j=1}^{N}2\pi K_j\sum_{l=1}^{\infty}\sin l\theta_j e^{-l^2\sigma_j^2/2}B_l(\omega) \tag{24}$$

where

$$B_l(\omega)=\frac{\hat{A}_{l-1}(\omega)-\hat{A}_{l+1}(\omega)}{2}\qquad\text{if } l\neq 1$$

$$B_1(\omega)=\hat{A}_0(\omega)-\frac{\hat{A}_2(\omega)}{2}$$

$$\hat{A}_0(\omega)=\frac{1}{2}\left\{\sum_{\mu=0}^{\infty}\{[\hat{a}_\mu(\omega)]_R^2+[\hat{a}_\mu(\omega)]_I^2\}+[\hat{a}_0(\omega)]_R^2+[\hat{a}_0(\omega)]_I^2\right\}$$

and

$$\hat{A}_k(\omega)=\sum_{v=0}^{\infty}\{[\hat{a}_{v+k}(\omega)]_R[\hat{a}_v(\omega)]_R+[\hat{a}_{v+k}(\omega)]_I[\hat{a}v(\omega)]_I\}+\frac{1}{2}\left\{\frac{[a_k(\omega)]_R^2}{2}+\frac{[a_k(\omega)]_I^2}{2}\right\}\left\{\frac{1+(-1)^k}{2}\right\}+\sum_{v=0}^{\left[\frac{k}{2}-1,\frac{k-1}{2}\right]}\{[\hat{a}_{k-v}(\omega)]_R[\hat{a}_v(\omega)]_R+[\hat{a}_{k-v}(\omega)]_I[\hat{a}_v(\omega)]_I\}$$

for $k>0$ where $\left[\frac{k}{2}-1, \frac{k-1}{2}\right]$ means $\frac{k}{2}-1$ if k is even, $\frac{k-1}{2}$ if k is odd, where the subscripts R and I refer to "Real part" and "Imaginary part" respectively, and where

$$[a_\mu(\omega)]_R=\sum_{n=0}^{\infty}(2n+1)[F_n(i\omega)]_R\tilde{a}_{n\mu}$$

$$[\hat{a}_\mu(\omega)]_I=\sum_{n=0}^{\infty}(2n+1)[F_n(i\omega)]_I\tilde{a}_{n\mu}$$

Normalization of $p(\theta)$ can be accomplished by integrating (18) subject to the above assumptions (1) and (2). The result is

$$2p_0+\sum_{j=1}^{N}K_j\sin\theta_j e^{-\sigma_j^2/2}=1. \tag{25}$$

The condition (25) will serve as a constraint on the relative values of p_0 and the Kj coefficients.

8. Results and Conclusions

Note that the theoretical resonant frequencies for a lossless ionosphere are given by the numbers $\frac{1}{(2\pi)}\sqrt{\left(\frac{c}{a}\right)^2 n(n+1)}$, which for $c/a=46.7$ are 10.5, 18.2, 25.6, 33.1, and 40.5. These are sufficiently different from the experimentally observed resonances to warrant the conclusion that ionospheric losses play a major role in the mode spectrum. The real positive quantity α, equal to $1/(h\sqrt{\mu\sigma_i})$ is a measure of these losses. To study its effect, the functions $\omega_n F_n(i\omega)$ were calculated for various values of α. In the earliest work, the mode series was computed for a few values of α (assumed independent of frequency) corresponding to $h\approx 80$ km (a compromise between day and night ionospheres) and values of σ_i between (10^{-8}) and (10^{-6})mhos/m. It was found that deviations from observed mode resonant frequencies were very substantial (e.g., 20 to 30 percent) and that no constant value of α would produce near-agreement on all five observed modes. The next approach was to allow α to vary with frequency and use the Balser and Wagner results to infer its frequency variation. This was initially done by graphical methods based on plots of real and imaginary parts of $\omega_n F_n(i\omega)$ [Raemer, 1961]. The method has been revised to increase accuracy. The function $|\omega_n F_n(i\omega)|^2$ was computed in terms of normalized variables $x_n\equiv\omega/\omega_n$ and $y_n\equiv\frac{\alpha}{\sqrt{2}\omega_n}$ (plots are shown in fig. 1) and its derivative with respect to x_n equated to zero to find the maximum. Corresponding values of y_n were then found for $n=1$ through 5, resulting in a function $\alpha(f)$ required to match theoretical and observed resonant frequencies. This function was then fitted to an analytical curve, by the method of least squares and (see fig. 2) a satisfactory fit was found to be

$$\alpha(f)=7.649-0.1863f+0.01362f^2-0.00006602f^3-0.00000388f^4 \tag{26}$$

Note that this differs from the function originally reported by the author [1961], i.e.,

$$\alpha(f)=6.44\log_{10}f-0.0131(f-8.0) \tag{27}$$

This is due to the difference in computational method, and the function (26) should be regarded as the more accurate of the two.

The fundamental assumption behind the mode-matching technique described above is that the observed modes are those arising from the individual nth order frequency response functions of the earth-air-ionosphere cavity and are not significantly influenced either by cross-terms between modes, overlap between modes, or the details of the lightning flash spectrum.

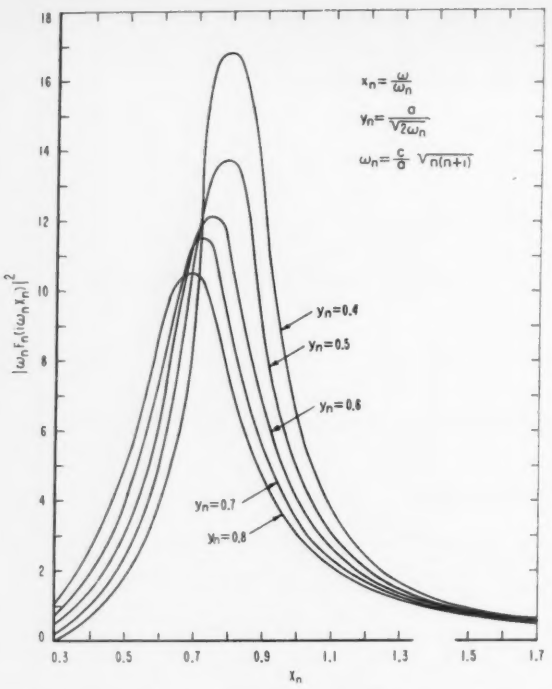

FIGURE 1. $|\omega_n F_n(i\omega_n x_n)|^2$ *versus* x_n *with* y_n *as parameter.*

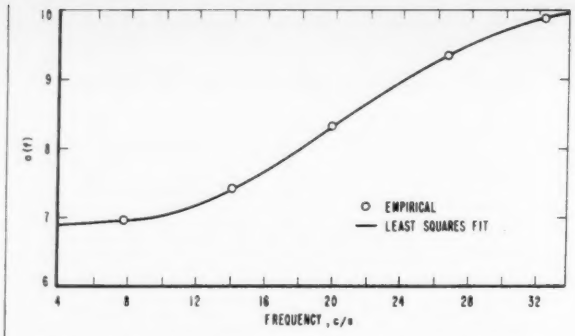

FIGURE 2. *Ionospheric loss parameter versus frequency.*

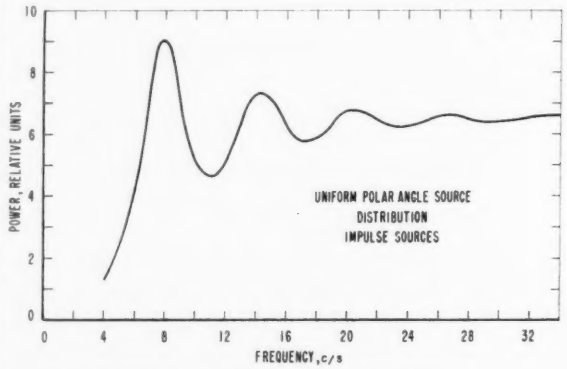

FIGURE 3. *Idealized ELF noise spectrum.*

For the moment we confine attention to the uniform angular distribution of sources. Using (26) to represent α, the quantity given by (19) has been calculated and is shown plotted in figure 3.

The plot of figure 3 is not appropriate for comparison with the experimental curve, since it assumes a white lightning flash spectrum and neglects the fact (see (1)) that the field amplitude contains a factor proportional to the reciprocal of the effective ionosphere height h. If it is assumed that the frequency dependence of α is due entirely to the effective ionosphere height and not to frequency dependence in the effective conductivity σ_i, then for consistency the power should contain a factor $[\alpha(f)]^2$. Multiplication of the quantity plotted in figure 3 by this factor leads to degradation rather than enhancement of agreement with experimental results, i.e., the envelope increases rapidly with frequency. Assuming our basic model to be essentially correct, one of two alternative possibilities follows: (1) The frequency dependence in α is entirely due to h and the spectrum of lightning flashes within the frequency range of interest is substantially different from a "white noise" spectrum or (2) the frequency dependence in α is largely due to σ_i, in which case the lightning spectrum still cannot be "white" noise but need not decay very rapidly with frequency within the range of interest.

Assuming that the ELF noise is largely caused by return lightning strokes, the suspicion that there is a significant decay in the spectrum with increasing frequency is borne out by figure 4. This figure, showing the estimated spectrum of a lightning flash compounded from (9) and (17) shows an average spectral decay of between 3 and 4 db per octave of frequency in the region of interest. Inserting this factor into (19) results in the theoretical curve (a) of figure 5. Incorporating the frequency dependence of h into the calculation results in the theoretical curve of figure 6. On both figures 5 and 6, the experimental curve is also shown for purposes of comparison. In both figures the experimental curve has been normalized in such a way that its average power is equal to the average power of the theoretical curve.

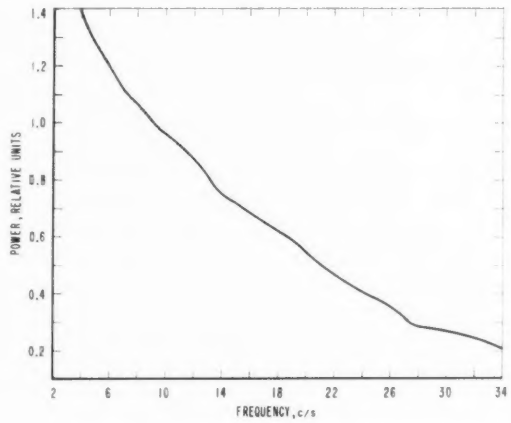

FIGURE 4. *Estimated lightning flash spectrum.*

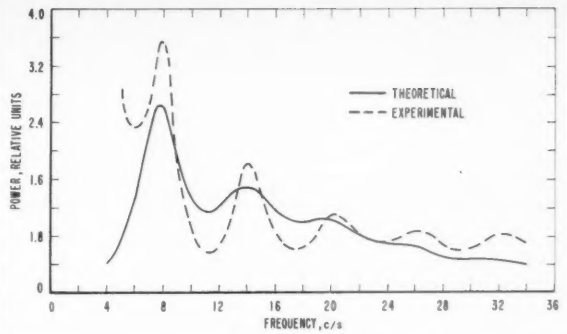

FIGURE 5. *ELF noise spectrum—effective ionospheric height assumed constant.*

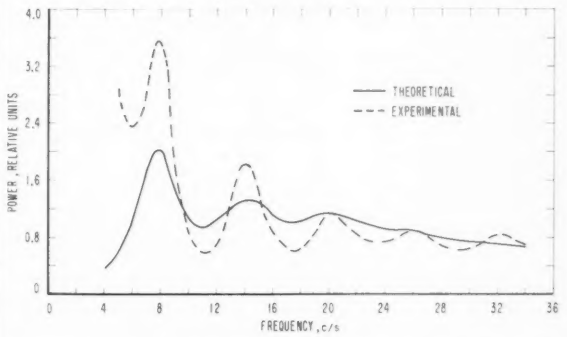

FIGURE 6. *ELF noise spectrum—effective ionospheric height assumed frequency dependent.*

The minima in the experimental curves are somewhat deeper than those of the theoretical spectra and the envelope of the minima, which is roughly constant on the experimental curves, decays substantially with frequency in the theoretical results. The rate of decay of the envelope of the maxima is comparable, but too rapid at the high frequency end of the region of interest. Widths of the modes are roughly comparable. Generally the agreement of spectral shape is seen to be reasonably good near the lower order modes and rather poor near the higher order modes. On the theoretical curves of both figures 5 and 6, the fourth mode becomes barely visible and the fifth is completely suppressed. Agreement is somewhat better in general in the case where the frequency dependence of $1/h^2$ is not accounted for, providing some evidence that the frequency dependence of α is predominantly due to the conductivity σ_t.

The computations on the equatorial region analysis are currently in progress. Preliminary results seem to indicate that accounting for these regions has no effect on the depths of the modes, but will have some influence on the envelope of the spectrum.

Other extensions which might reduce the discrepancies between the theory presented here and the Balser and Wagner results are as follows:

(1) The lightning statistics used here were obtained from data taken at many times of year and at many locations. Long term (e.g., year-round) averages of lightning parameters are not necessarily the same parameters averaged over a single day. The ELF noise results obtained by Balser and Wagner [1960] were obtained on a single day and at a single location. Only observed spectra averaged over many locations and times of year would be strictly comparable with the theory discussed here, unless the ELF noise spectrum is found to be essentially independent of time and location, a supposition that would require many more experiments to ascertain.

(2) The same remarks apply to the assumption that the lightning which gives rise to ELF noise consists entirely of cloud-ground return strokes and that the lightning current is entirely vertical. Wait [1960a] has pointed out that there may be significant contributions from horizontal current components at the frequencies of interest here.

(3) Assuming the return lightning stroke moment to be a random function and averaging over all random variables would probably make the lightning stroke more "noisy," i.e., give it a flatter spectrum. This could result in a theoretical spectrum whose envelope bears a closer resemblance to that of the observed spectrum.

(4) The ionosphere model could be improved by accounting for the earth's magnetic field and the effects of heavy ions.

(5) The probable reason for the fact that our theoretical curves are not as oscillatory as the experimental results is that the assumed ionospheric loss parameter is too large, tending to "flatten" the modes. This implies that the homogeneous ionsophere model is not sufficiently realistic for precise calculation of ELF noise effects. In a model accounting for the variation of conductivity with height near the lower edge of the ionosphere, e.g., the exponential model of J. R. Wait [1960b, 1960c], the loss parameter would be a complex number.[7] This would result in a more oscillatory spectral curve, as indicated in recent work by J. Galejs [private communication], as yet unpublished, and in some preliminary calculations by the author. Work by the author to infer information about the ionosphere from the Balser and Wagner curves with the assumption of a complex loss parameter is currently in progress.

Appreciation is expressed to Dr. R. Row and Dr. J. Galejs for helpful discussions pertaining to the work reported here, and to Mrs. J. Van Horn and Mr. C. Roche for computational assistance.

9. Appendix. Remarks on Analytical Points

Note that with a uniform distribution, a source could be at the observer's position. The discontinuity of the field at this point is accounted for in the fundamental theory by the fact that the hypergeometric series $P_\nu(-\cos\theta)$ used in the representation of the field has a discontinuity at $\theta=0$. The expansion of this function into a series of Legendre polynomials [Magnus and Oberhetinger, 1949], which is basic to our formulation, is valid everywhere but at $\theta=0$. Sommerfeld [1949] has shown that $P_\nu(-\cos\theta)$ becomes proportional to $\log\theta^2$ as θ approaches zero. In carrying out the integration indicated in (7) in the vicinity of this point, the product $(\sin\theta \log\theta^2)$ appearing in the integrand is an indeterminate form which can be shown to vanish as $\theta\to 0$. The contribution to the integral at $\theta=0$ therefore vanishes, and no difficulties are encountered in extending the expansion to this point.

Another analytical point worthy of note is that the expansion mentioned above is not valid at $\nu=0, \pm1, \pm2, \pm3, \ldots$ etc. The point $\nu=0$ occurs at zero frequency, which we exclude from the analysis. Also, it can be shown that, except at $\omega=0$, ν always has an imaginary part and therefore cannot take on values $\pm1, \pm2, \ldots$ etc.

10. References

Balser, M., and C. A. Wagner, Observations of earth-ionosphere cavity resonances, Nature **188**, pp. 638–641 (Nov. 19, 1960).

Davenport, W. B. and W. L. Root, Random signals and noise, ch. 7 (McGraw-Hill Book Co., Inc., New York, N.Y., 1958).

Handbook of Geophysics, U.S. Air Force, ch. 9 (MacMillan & Co., New York, N.Y., 1960).

Jahnke-Emde, Tables of functions, p. 107 (Dover Publs., New York, N.Y., 1945).

Magnus, W., and F. Oberhetinger, Special functions of mathematical physics, p. 57 (Chelsea Publ. Co., New York, N.Y., 1949).

---

[7] This would account for both the ionosphere attenuation loss undergone by a propagating wave and a phase shift due to the conductivity gradient, whereas the present treatment accounts only for the former. Neglect of the phase shift would be valid with a sharply bounded homogeneous ionosphere.

Raemer, H. R., On the extremely low frequency spectrum of earth-ionosphere cavity response to electrical storms, J. Geophys. Research **66,** No. 5, 1580–1583 (May 1961).
Schumann, W. O., Uber electrische Eigenschwingungen des Hohlraumes Erde-Luft Ionosphäre, erregt durch Blitzentladungen, Zeitschrift für angewandte Physik **8,** pp. 373–378 (August 1957).
Sommerfeld, A. N., Partial differential equations, p. 215 (Academic Press, New York, N.Y., 1949).
Stratton, J. A., Electromagnetic theory, pp. 326–327 (McGraw-Hill Book Co., Inc., New York, N.Y., 1941).
Wait, J. R., Mode theory and the propagation of ELF radio waves, J. Research NBS (Radio Prop.) **64D,** No. 4, 387–404 (July–August 1960a).
Wait, J. R., Terrestrial propagation of very-low-frequency radio waves, J. Research NBS (Radio Prop.) **64D,** No. 2, 153–204 (March–April 1960b).
Wait, J. R., On the propagation of ELF radio waves and the influence of a nonhomogeneous ionosphere, J. Geophys. Research **65,** pp. 587–607 (Feb. 1960c).
Wax, N. (Editor), Papers on noise and stochastic processes, sec. 1, p. 145, Rice, S. O., Mathematical analysis of random noise (Dover Publs., New York, N.Y., 1954).
Williams, J. C., Thunderstorms and VLF radio noise, Ph. D. Thesis, Harvard University, Division of Engineering and Applied Physics (May 1959).

(Paper 65D6–162)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

The Nonsingular Embedding of Transition Processes Within a More General Framework of Coupled Variables

John Heading

Contribution from the University of Southampton, Southampton, England

(Received June 1, 1961)

Reflection and coupling processes exhibited by plane electromagnetic waves propagated in an inhomogeneous horizontally-stratified anisotropic ionosphere are associated with discrete transition points or with continuous coupling regions. These arise when the fourth order differential equations are written in first order coupled form, and many terms in these equations become infinite at the transition points. This procedure is rendered more precise by means of a special linear transformation that reformulates the equations in a new way, thereby exhibiting the manner in which local coupling processes are embedded in the more general background process of otherwise independently propagated characteristic waves. To exhibit the power of the matrix algebra involved, the case of an arbitrary number of characteristic waves is considered; moreover, Försterling-type coupled equations are produced in a more generalized form than hitherto considered, and a discussion of the equations governing continuous coupling completes the paper.

1. Introduction

The equations governing the propagation of plane electromagnetic waves in a horizontally-stratified inhomogeneous anisotropic ionosphere may be expressed in various ways, depending upon the particular investigation for which the equations are required. If the Oz-axis of a right-handed set of Cartesian axes Ox, Oy, Oz is directed vertically upwards, then it often proves convenient to use second or fourth order differential equations in the dependent variables E_x or E_y. Heading and Whipple [1952] and Heading [1955] have employed this representation when analytical solutions of the differential equations are under consideration. The reason is that all known analytical solutions of the field equations may be expressed in terms of generalized hypergeometric functions, hypergeometric functions, confluent hypergeometric functions, Bessel functions and so on, these functions naturally satisfying second or fourth order differential equations. Such functions provide the reflection and conversion coefficients for a particular anisotropic model ionosphere, but by inspecting the functional form of the electric field throughout the medium no immediate physical interpretation is possible, whereby the actual process of wavereflection and wavecoupling may be explicitly exhibited.

Sets of equations known as *coupled equations* have therefore been introduced, in order to enable approximate solutions for the electromagnetic field to be obtained, and in order to exhibit directly the important physical processes of wave reflection and wave coupling. Försterling [1942] considered a coupled system of two second order equations; these refer to an ionospheric model in which the external magnetic field is oblique but in which propagation is vertical. Gibbons and Nertney [1951, 1952] have used these coupled equations to obtain approximate numerical solutions for various ionospheric models, while Budden [1952] has used them in discussing limiting polarization. Heading [1953] and Budden and Clemmow [1957] have given similar coupled equations governing the case when incidence is oblique upon an ionosphere in which the magnetic field is vertical.

When both the direction of propagation and of the magnetic field are oblique, four coupled equations are used, each of the first order. These were first introduced by Bremmer [1949], but more systematically using matric techniques by Clemmow and Heading [1954]. The most up-to-date account of these first order coupled equations is found in chapter 18 of the recent text by Budden [1961], but there the fundamental simplicity and beauty of the matrix formula-

tion is often disguised under a multitude of symbols and approximations. Moreover, certain important features of the matrices occurring in the coupled equations together with the possibility of nonsingular embedding do not as yet appear to have been recognized.

The coupled equations exhibit those ranges of height in which the propagation of the four characteristic waves takes place almost independently. Any point at which many of the coefficients in the equations become singular denotes that coupling takes place there between the characteristic waves concerned, the other characteristic waves not embraced by the coupling still being propagated independently.

The present investigation consists of an examination of the equations near the coupling points without the necessity of introducing approximations. It will be proved that the equations governing the coupling process at a coupling point are embedded in the equations governing the propagation of the waves as a whole, such that all singularities in the coefficients of all equations are removed at the coupling point. Moreover, although ionospheric propagation is represented in terms of four first order equations, yielding four characteristic waves in a homogeneous or slowly-varying medium, generalization to wave processes governed by n first order simultaneous differential equations yielding n characteristic waves is easily possible using matrix notation. The sphere of usefulness of the present investigation will thereby be extended if n equations are considered.

2. Transformation of an rth Order Linear Equation

The preliminary study of a given linear differential equation rather than a set of first order simultaneous linear differential equations is necessary in order to define a *principal embedded coupled matrix*.

The rth order linear equation in the dependent variable u

$$u^{(r)}=v_{r-1}(z)u^{(r-1)}+v_{r-2}(z)u^{(r-2)}+\ldots+v_1(z)u'+v_0(z)u, \tag{1}$$

where bracketed superscripts denote differentiation with respect to z, and where the r coefficients $v_0, v_1, \ldots, v_{r-1}$ are given functions of z, may be expressed as a linear system using matrix notation as follows:

$$\frac{d}{dz}\begin{bmatrix} u \\ u' \\ \cdots \\ u^{(r-2)} \\ u^{(r-1)} \end{bmatrix}=\begin{bmatrix} 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 & 0 \\ \multicolumn{7}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \\ v_0 & v_1 & v_2 & v_3 & \ldots & v_{r-2} & v_{r-1} \end{bmatrix}\begin{bmatrix} u \\ u' \\ \cdots \\ u^{(r-2)} \\ u^{(r-1)} \end{bmatrix}, \tag{2}$$

or

$$\mathbf{u}'=\mathbf{T}\,\mathbf{u} \tag{3}$$

say, where $\mathbf{u}$ denotes the column consisting of r elements and $\mathbf{T}$ denotes the square $r\times r$ matrix shown consisting of the coefficients occurring in (1).

When the matrix $\mathbf{T}$ is diagonalized by means of the transformation

$$\mathbf{u}=\mathbf{A}\,\mathbf{f},$$

the roots of the characteristic equation

$$\det(\mathbf{T}-q\mathbf{I})=0$$

are required. This equation has the explicit form

$$q^r - v_{r-1}q^{r-1} - v_{r-2}q^{r-2} - \ldots - v_1 q - v_0 = 0,$$

possessing the r characteristic roots $q_1, q_2, \ldots, q_r$.

The elements of the characteristic vector corresponding to the root q_j consist of the cofactors of any row of the matrix $\mathbf{T} - q_j\mathbf{I}$; in particular, the cofactors of the last row are $1, q_j, q_j^2, \ldots, q_j^{r-1}$ respectively. Hence the matrix $\mathbf{A}$ of the transformation may be taken to be

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ q_1 & q_2 & \ldots & q_r \\ \ldots & \ldots & \ldots & \ldots \\ q_1^{r-1} & q_2^{r-1} & \ldots & q_r^{r-1} \end{bmatrix};$$

this is an *alternant* matrix, whose determinant is given by

$$\det \mathbf{A} = \prod_{\substack{i,j=1 \\ i>j}}^{r} (q_i - q_j). \tag{4}$$

It follows that $\mathbf{A}$ is nonsingular throughout any domain in the complex z-plane in which all the r characteristic roots are distinct.

Equation (3) becomes

$$\mathbf{A'f} + \mathbf{Af'} = \mathbf{TAf},$$

or

$$\mathbf{f'} = \mathbf{A^{-1}TAf} - \mathbf{A^{-1}A'f}.$$

The matrix $\mathbf{A^{-1}\,T\,A}$ is now a diagonal matrix $\mathbf{Q}$, whose diagonal elements consist of the r characteristic roots in order; hence

$$\mathbf{f'} = \mathbf{Qf} - \mathbf{A^{-1}A'f}. \tag{5}$$

If solutions of (5) are known, then $\mathbf{u} = \mathbf{A\,f}$ are the solutions of (3), and in particular the solution of (1) is

$$u = f_1 + f_2 + \ldots + f_r. \tag{6}$$

The r eqs (5) constitute a set of first order *coupled* equations equivalent to the original eq (1). The diagonal matrix $\mathbf{Q}$ will be called a *principal uncoupled matrix*, while the terms $\mathbf{A^{-1}\,A'\,f}$ will be called the *coupling terms*, though strictly speaking coupling arises only through the nondiagonal elements of the matrix $\mathbf{A^{-1}A'}$. This product $\mathbf{A^{-1}A'}$, whose explicit elements are calculated in section 5, may be called the *coupling matrix*. It is the presence of this coupling matrix that renders the r eqs (5) simultaneous and not independent, but the coupling matrix vanishes if the coefficients in the original eq (1) are constants, owing to the presence of derivatives in every element. The r transformed equations are then completely independent, and their r solutions provide r characteristic waves independently propagated when a physical picture is employed for their interpretation. Then

$$f_j = C_j \exp(q_j z),$$

where the C_j are arbitrary constants of integration.

When the coefficients in (1) are not constants, usually some parameter such as the frequency or some property of the medium attains values that cause the elements of the coupling matrix

to be small in magnitude compared with the elements of the principal uncoupled matrix, provided a domain is considered in which **A** is nonsingular so that its reciprocal exists. Under these circumstances, the nondiagonal elements in the coupling matrix are neglected, yielding r independent equations, whose solutions represent r characteristic waves propagated independently through the slowly varying medium.

In this paper, equations of the form (1) are considered only when all the r characteristic roots become equal at a particular point z_0. To this end, r new functions $\alpha, \beta, \ldots, \rho$ are introduced, defined by the r linear equations

$$q_j=\alpha+k^{j-1}\beta+k^{2(j-1)}\gamma+\ldots+k^{(r-1)(j-1)}\rho, \tag{7}$$

where $k=\exp(2\pi i/r)$. The matrix of the coefficients in these equations is nonsingular, since it is an alternant whose elements consist of the rth roots of unity 1, k, k^2, $\ldots$, k^{r-1} raised to the appropriate powers. The r new functions are therefore uniquely defined.

Now

$$v_{r-1}=\sum_{j=1}^{r} q_j$$

$$=r\alpha+\beta\sum_{j=1}^{r} k^{j-1}+\gamma\sum_{j=1}^{r} k^{2(j-1)}+\ldots+\rho\sum_{j=1}^{r} k^{(r-1)(j-1)}.$$

But if p is an integer,

$$\left[\sum_{j=1}^{r} k^{p(j-1)}\right](1-k^p)=1-(k^p)^r=1-(k^r)^p=0,$$

so if $p\neq 0$, it follows that

$$\sum_{j=1}^{r} k^{p(j-1)}=0. \tag{8}$$

Hence

$$\alpha=v_{r-1}/r.$$

For example, when $r=2$, α and β are merely the two terms that arise when the quadratic equation is solved using the standard formula, the familiar $\pm$ signs being associated with the β. When $r=3$, β and γ arise when the reduced cubic is solved by Cardan's method.

To lay the basis for the following sections, v_{r-1} is now chosen to vanish identically, indicating that $\alpha=0$. It will be seen in the general theory, however, that no loss of generality occurs on account of this assumption. Coupling is then studied by postulating that the r characteristic roots become equal at $z=z_0$. Under these circumstances, when α is omitted from (7) and when the r q's are given equal non-zero values, the rank of the matrix of the coefficients of the right-hand side is $r-1$, since the r linear expressions are dependent, while the rank of the augmented matrix is r. The equations are therefore inconsistent. If on the other hand the r equal values of the q's are zero, the first $r-1$ equations (7) yield $\beta=\gamma=\ldots=\rho=0$, since the matrix of the coefficients of these first $r-1$ equations is nonsingular.

Hence, at $z=z_0$, it follows that

$$q_1=q_2=\ldots=q_r=0;$$

moreover, the coefficients of (1) in which $v_{r-1}\equiv 0$ consist of sums of products of these r roots, so

$$v_0(z_0)=v_1(z_0)=\ldots=v_{r-2}(z_0)=0.$$

This is too general for useful progress to be made, although the theory to be presented is equally valid without any further simplifying assumptions. The special choice is made that

$$v_1(z) \equiv v_2(z) \equiv \ldots \equiv v_{r-2}(z) \equiv 0,$$

leaving a differential equation

$$\frac{d^r u}{dz^r} = v_0(z) u. \tag{9}$$

When $r=2$, it should be noticed that no assumption is in fact made, and it is just this case that usually governs wave reflection and wave coupling. The r values of the characteristic roots are now given by

$$q_j = k^{j-1}(v_0)^{1/r} \equiv k^{j-1}\beta.$$

Another reason why these simplifying assumptions are advantageous concerns the possibility of writing down approximate solutions of (9), since (9) will form what may be termed a *comparison* equation embedded in an equation of higher order. Approximate solutions of this latter general equation are given in terms of approximate solutions of the comparison equation. When $r=2$, (9) can be solved approximately in terms of the Airy integral, while if $r > 2$, solutions may be obtained in terms of functions investigated by Heading in a series of papers [1957 a and b; 1960], though such approximations have not as yet been published. But if (1) is considered with $v_{r-1} \equiv 0$ but with no further assumptions, useful approximate analytical solutions are almost impossible to obtain near a point $z=z_0$ at which r characteristic roots become equal.

Under these assumptions, matrix **T** becomes

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \\ v & 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix} \tag{10}$$

where v_0 is now written simply as v.

The characteristic equation $q^r - v = 0$ has the r roots

$$q_j = k^{j-1} v^{1/r}$$

and the alternant matrix **A** factorizes thus:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & v^{1/r} & 0 & \ldots & 0 \\ 0 & 0 & v^{2/r} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & v^{(r-1)/r} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & k & k^2 & \ldots & k^{r-1} \\ 1 & k^2 & k^4 & \ldots & k^{2(r-1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & k^{r-1} & k^{2(r-1)} & \ldots & k^{(r-1)^2} \end{pmatrix} \tag{11}$$

$$= \mathbf{VK}$$

say, where **V** and **K** are defined by these two square matrices.

The matrix $\mathbf{K}$ being constant, it follows that

$$\mathbf{A}^{-1}\mathbf{A}'=\mathbf{K}^{-1}\mathbf{V}^{-1}\mathbf{V}'\mathbf{K}$$

$$=\mathbf{K}^{-1}\begin{pmatrix}1 & 0 & 0 & \dots\\ 0 & v^{-1/r} & 0 & \dots\\ 0 & 0 & v^{-2/r} & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\begin{pmatrix}0 & 0 & 0 & \dots\\ 0 & v^{1/r}v'/rv & 0 & \dots\\ 0 & 0 & 2v^{2/r}v'/rv & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\mathbf{K}$$

$$=\frac{v'}{rv}\mathbf{K}^{-1}\begin{pmatrix}0 & 0 & 0 & \dots\\ 0 & 1 & 0 & \dots\\ 0 & 0 & 2 & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\mathbf{K}.$$

Clearly coupling only becomes large near a point at which v'/v is singular.

The explicit elements of this product have been calculated by Heading [1960]. They are

$$(\mathbf{A}^{-1}\mathbf{A}')_{jj}=\frac{v'}{v}\cdot\frac{r-1}{2r}, \tag{12}$$

$$(\mathbf{A}^{-1}\mathbf{A}')_{ij}=\frac{v'}{v}\cdot\frac{1}{r(k^{j-i}-1)}. \tag{13}$$

The coupled eqs (5) then take the form

$$=\begin{pmatrix}v^{1/r} & 0 & 0 & \dots\\ 0 & kv^{1/r} & 0 & \dots\\ 0 & 0 & k^2v^{1/r} & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\mathbf{f}-\frac{v'}{v}\cdot\frac{r-1}{2r}\mathbf{f}-\frac{v'}{rv}\begin{pmatrix}0 & \frac{1}{k-1} & \frac{1}{k^2-1} & \dots\\ \frac{1}{k^{-1}-1} & 0 & \frac{1}{k-1} & \dots\\ \frac{1}{k^{-2}-1} & \frac{1}{k^{-2}-1} & 0 & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\mathbf{f}. \tag{14}$$

The first square matrix on the right hand side of (14) is the principal uncoupled matrix. The coupling terms become singular whenever v'/v is singular; such a point is known as a reflection or coupling point. Mathematically speaking, such a point is called a *transition point*. The coupling terms are singular at this point, but the original equation

$$\mathbf{u}'=\mathbf{T}\mathbf{u}$$

possesses a special matrix (10), associated in this particular case with no additional coupling terms. An $r\times r$ matrix of the form (10) will be called a *principal embedded coupled matrix*.

3. General System

The equations governing ionospheric radio propagation may be expressed as four linear first order differential equations with four dependent variables, given for example by Clemmow and Heading [1954]. More generally, consider n dependent variables $e_1, e_2, \dots, e_n$ satisfying the n linear first order equations

$$e_1' = \sum_{j=1}^{n} T_{ij} e_j.$$

In maxtrix notation,

$$\mathbf{e}' = \mathbf{T}\mathbf{e}, \tag{15}$$

where the $n \times n$ matrix $\mathbf{T}$ contains n^2 elements each of which is a given function of z. In the ionospheric case, $n=4$, and some of the elements of $\mathbf{T}$ are zero, while others are complicated functions of height, the explicit forms of which would obscure the essential simplicity of the present argument. The matrix $\mathbf{T}$ is therefore more complicated than that occurring in (3) or the principal embedded coupled matrix (10).

If the dependent column $\mathbf{e}$ is now transformed to the column $\mathbf{f}$ by the transformation

$$\mathbf{e} = \mathbf{S}\mathbf{f}, \tag{16}$$

it follows that

$$\mathbf{f}' = \mathbf{S}^{-1}\mathbf{T}\mathbf{S}\mathbf{f} - \mathbf{S}^{-1}\mathbf{S}'\mathbf{f}.$$

If the n characteristic roots of $\mathbf{T}$ are denoted by $q_1, q_2, \ldots, q_n$, let $\mathbf{s}_j$ be the column matrix formed from the cofactors of $\mathbf{T} - q_j\mathbf{I}$ taken along any suitable row. Then the transformation matrix

$$\mathbf{S} = (\mathbf{s}_1 \mathbf{s}_2 \ldots \mathbf{s}_n)$$

diagonalizes the matrix $\mathbf{T}$; that is, $\mathbf{S}^{-1}\mathbf{T}\,\mathbf{S} = \mathbf{Q}$ consists of the n roots arranged in order down the leading diagonal. The coupled equations become

$$\begin{aligned} \mathbf{f}' &= \mathbf{Q}\mathbf{f} - (\mathbf{s}_1 \mathbf{s}_2 \ldots \mathbf{s}_n)^{-1}(\mathbf{s}_1 \mathbf{s}_2 \ldots \mathbf{s}_n)'\mathbf{f} \\ &= \mathbf{Q}\mathbf{f} - \mathbf{S}^{-1}\mathbf{S}'\mathbf{f}, \end{aligned} \tag{17}$$

valid at all points at which $\mathbf{S}$ is nonsingular.

If desired, a change of independent variable may easily be effected. If $\zeta = \zeta(z)$, then

$$\frac{d\mathbf{f}}{dz} = \frac{d\mathbf{f}}{d\zeta}\zeta',$$

yielding

$$\frac{d\mathbf{f}}{d\zeta} = \frac{1}{\zeta'}\mathbf{Q}\mathbf{f} - \mathbf{S}^{-1}\frac{d\mathbf{S}}{d\zeta}\mathbf{f}.$$

Similarly, a change of dependent variables may be made by placing

$$f_j = \phi_j(z) g_j,$$

where the n functions $\phi_j(z)$ denotes n given functions. In matrix form, let

$$\mathbf{f} = \mathbf{\Phi}\mathbf{g},$$

where $\mathbf{\Phi}$ is a diagonal matrix consisting of the n functions $\phi_j(z)$ arranged in order down the leading diagonal. Then

$$\mathbf{\Phi}\mathbf{g}' + \mathbf{\Phi}'\mathbf{g} = \mathbf{Q}\mathbf{\Phi}\mathbf{g} - \mathbf{S}^{-1}\mathbf{S}'\mathbf{\Phi}\mathbf{g},$$

or

$$\begin{aligned} \mathbf{g}' &= \mathbf{\Phi}^{-1}\mathbf{Q}\mathbf{\Phi}\mathbf{g} - (\mathbf{\Phi}^{-1}\mathbf{\Phi}' + \mathbf{\Phi}^{-1}\mathbf{S}^{-1}\mathbf{S}'\mathbf{\Phi})\mathbf{g} \\ &= \mathbf{Q}\mathbf{g} - (\mathbf{S}\mathbf{\Phi})^{-1}(\mathbf{S}\mathbf{\Phi})'\mathbf{g}, \end{aligned}$$

implying that these equations would have been produced had $\mathbf{S}\mathbf{\Phi}$ been used for the transformation matrix (16) rather than $\mathbf{S}$. The characteristic vector $\mathbf{s}_j$ is merely multiplied throughout by ϕ_j.

Now each cofactor of $\mathbf{T}-q_j\mathbf{I}$ is a polynomial of degree $n-1$ at most in q_j, the coefficients being independent of the suffix j. Generally,

$$\mathbf{s}_j=\begin{pmatrix}\text{polynomial 1 of degree } n-1 \text{ at most in } q_j\\ \text{polynomial 2 of degree } n-1 \text{ at most in } q_j\\ \cdots\cdots\cdots\cdots\\ \text{polynomial } n \text{ of degree } n-1 \text{ at most in } q_j\end{pmatrix}$$

$$=\begin{pmatrix}\text{the } n \text{ coefficients of polynomial 1}\\ \text{the } n \text{ coefficients of polynomial 2}\\ \cdots\cdots\cdots\cdots\\ \text{the } n \text{ coefficients of polynomial } n\end{pmatrix}\begin{pmatrix}1\\ q_j\\ \cdots\\ q_j^{n-1}\end{pmatrix}$$

$$=\mathbf{P}\begin{pmatrix}1\\ q_j\\ \cdots\\ q_j^{n-1}\end{pmatrix}$$

say, where $\mathbf{P}$ denotes the $n\times n$ matrix formed from the coefficients of the n polynomials. The last column of $\mathbf{P}$ consists of $n-1$ zero elements, while the remaining elements is unity (when n is even), since only one polynomial can be of degree $n-1$, the others being of degree $n-2$.

Then

$$\mathbf{S}=(\mathbf{s}_1\mathbf{s}_2\ldots\mathbf{s}_n)=\mathbf{P}\begin{pmatrix}1 & 1 & 1 & \ldots\\ q_1 & q_2 & q_3 & \ldots\\ q_1^2 & q_2^2 & q_3^2 & \ldots\\ \cdots & \cdots & \cdots & \cdots\end{pmatrix}=\mathbf{PA},$$

where $\mathbf{A}$ denotes the alternant matrix consisting of the roots and their powers arranged in order. It follows that

$$\mathbf{S}^{-1}\mathbf{S}'=\mathbf{A}^{-1}\mathbf{P}^{-1}(\mathbf{PA}'+\mathbf{P}'\mathbf{A})$$

$$=\mathbf{A}^{-1}\mathbf{A}'+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}.$$

The coupled equations (17) take the form

$$\mathbf{f}'=\mathbf{Qf}-\mathbf{A}^{-1}\mathbf{A}'\mathbf{f}-\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{Af},$$

similar to the set (5) apart from the additional coupling terms $\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}\mathbf{f}$. As before, $\mathbf{Q}$ denotes the principal uncoupled matrix, $\mathbf{A}^{-1}\mathbf{A}'$ is the *primary coupling matrix* and $\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}$ is the *secondary coupling matrix*.

4. Integral Equation

Let the matrix $\mathbf{A}^{-1}\mathbf{A}'$ be separated into the sum of two parts, $\mathbf{D}$ and $(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D})$ say, where $\mathbf{D}$ is an arbitrary diagonal matrix, though this may preferably consist of certain diagonal terms chosen from the explicit representation of $\mathbf{A}^{-1}\mathbf{A}'$ given in the following section. Then

$$\mathbf{f}'=(\mathbf{Q}-\mathbf{D})\mathbf{f}-(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f},$$

where the matrix $\mathbf{Q}-\mathbf{D}$ is diagonal.

Let $\mathbf{f}_0$ be a solution of the equation

$$\mathbf{f}'=(\mathbf{Q}-\mathbf{D})\mathbf{f},$$

this equation representing in effect n independent linear equations. Let $\mathbf{F}_0$ denote the $n\times n$ diagonal matrix containing the elements of $\mathbf{f}_0$ arranged down the leading diagonal. Consider the integral equation

$$\mathbf{f}=\mathbf{N}\mathbf{f}_0-\mathbf{F}_0\int_a^z \mathbf{F}_0^{-1}(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f}\,dz, \tag{19}$$

where $\mathbf{N}$ denotes an arbitrary constant diagonal matrix and a an arbitrary constant.

This is the required integral equation, for differentiation yields

$$\begin{aligned}\mathbf{f}'&=\mathbf{N}\mathbf{f}_0'-\mathbf{F}_0'\int_a^z \mathbf{F}_0^{-1}(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f}\,dz-\mathbf{F}_0\mathbf{F}_0^{-1}(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f}\\&=\mathbf{N}\mathbf{f}_0'-\mathbf{F}_0'\mathbf{F}_0^{-1}(\mathbf{N}\mathbf{f}_0-\mathbf{f})-(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f}\end{aligned}$$

from (19).

But

$$\mathbf{F}_0'=(\mathbf{Q}-\mathbf{D})\mathbf{F}_0$$

since all matrices are diagonal, so

$$\begin{aligned}\mathbf{f}'&=\mathbf{N}(\mathbf{Q}-\mathbf{D})\mathbf{f}_0-(\mathbf{Q}-\mathbf{D})(\mathbf{N}\mathbf{f}_0-\mathbf{f})-(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f}\\&=(\mathbf{Q}-\mathbf{D})\mathbf{f}-(\mathbf{A}^{-1}\mathbf{A}'-\mathbf{D}+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})\mathbf{f},\end{aligned}$$

since $\mathbf{N}$ commutes with $(\mathbf{Q}-\mathbf{D})$. This demonstrates the equivalence of the differential and integral equations.

5. Explicit Form of the Primary Coupling Matrix

In order to appreciate more clearly the character of the coupling terms $\mathbf{A}^{-1}\mathbf{A}'\mathbf{f}$, it is necessary to evaluate the explicit form of the product $\mathbf{A}^{-1}\mathbf{A}'$.

The reciprocal of the nth order alternant matrix

$$\mathbf{A}=\begin{pmatrix}1 & 1 & 1 & \dots & 1\\ q_1 & q_2 & q_3 & \dots & q_n\\ \dots & \dots & \dots & \dots & \dots\\ q_1^{n-1} & q_2^{n-1} & q_3^{n-1} & \dots & q_n^{n-1}\end{pmatrix}$$

may be calculated to have the form

$$\begin{pmatrix}\dfrac{(-1)^{n-1}S_{n-1}(q_1=0)}{(q_1)} & \dfrac{(-1)^{n-2}S_{n-2}(q_1=0)}{(q_1)} & \dots & \dfrac{-S_1(q_1=0)}{(q_1)} & \dfrac{S_0}{(q_1)}\\ \dfrac{(-1)^{n-1}S_{n-1}(q_2=0)}{(q_2)} & \dfrac{(-1)^{n-2}S_{n-2}(q_2=0)}{(q_2)} & \dots & \dfrac{-S_1(q_2=0)}{(q_2)} & \dfrac{S_0}{(q_2)}\\ \dots & \dots & \dots & \dots & \dots\\ \dfrac{(-1)^{n-1}S_{n-1}(q_n=0)}{(q_n)} & \dfrac{(-1)^{n-2}S_{n-2}(q_n=0)}{(q_n)} & \dots & \dfrac{-S_1(q_n=0)}{(q_n)} & \dfrac{S_0}{(q_n)}\end{pmatrix}.$$

Here, $S_0(\equiv 1)$, S_1, S_2, . . ., S_{n-1} are the elementary symmetric functions of the n quantities $q_1, q_2, \ldots, q_n$. The symbol $S_p(q_s=0)$ is used to denote the algebraic expression S_p when q_s is replaced by zero. The symbol (q_p) is used to denote the product

$$(q_p)\equiv(q_p-q_1)q_p-q_2)\ldots(q_p-q_{p-1})(q_p-q_{p+1})\ldots(q_p-q_n)=(\partial F/\partial q)_{q=q_p}, \tag{20}$$

where

$$F=\det(q\mathbf{I}-\mathbf{T}).$$

It can then be seen that

$$(\mathbf{A}^{-1})_{ij}=(-1)^{n-j}S_{n-j}(q_i=0)/(q_i),$$

and

$$(\mathbf{A}')_{ij}=(i-1)q_j^{i-2}q_j',$$

so the general element in the product $\mathbf{A}^{-1}\mathbf{A}'$ is given by

$$\begin{aligned}(\mathbf{A}^{-1}\mathbf{A}')_{ij}&=\sum_{k=1}^{n}(\mathbf{A}^{-1})_{ik}(\mathbf{A}')_{kj}\\&=\sum_{k=1}^{n}(-1)^{n-k}S_{n-k}(q_i=0)(k-1)q_j^{k-2}q_j'/(q_i).\end{aligned} \tag{21}$$

In order to simplify this expression, the polynomial of degree $n-1$ is formed, whose zeros are $q_1, q_2, \ldots, q_n$ but with q_i omitted.

This is

$$\begin{aligned}\frac{F}{q-q_j}&=q^{n-1}-S_1(q_i=0)q^{n-2}+\ldots+(-1)^pS_p(q_i=0)q^{n-p-1}+\ldots\\&=\sum_{p=0}^{n-1}(-1)^pS_p(q_i=0)q^{n-p-1}\\&=\sum_{k=1}^{n}(-1)^{n-k}S_{n-k}(q_i=0)q^{k-1},\qquad(\text{with } p=n-k)\end{aligned}$$

yielding upon differentiation with respect to q

$$\frac{\partial}{\partial_q}\left(\frac{F}{q-q_i}\right)=\sum_{k=1}^{n}(-1)^{n-k}S_{n-k}(q_i=0)(k-1)q^{k-2}. \tag{22}$$

If $i\neq j$, this result transforms (21) thus:

$$\begin{aligned}(\mathbf{A}^{-1}\mathbf{A}')_{ij}&=\frac{q_j'}{(q_i)}\left[\frac{\partial}{\partial q}\left(\frac{F}{q-q_i}\right)\right]_{q=q_j}\\&=\frac{q_j'}{(q_i)}\left[\frac{\partial F/\partial q}{q-q_i}-\frac{F}{(q-q_i)^2}\right]_{q=q_j}\\&=\frac{q_j'}{q_j-q_i}\cdot\frac{(q_j)}{(q_i)},\end{aligned} \tag{23}$$

since F vanishes when $q=q_j$. Out of the $n-1$ factors in (q_j), the factor (q_j-q_i) obviously cancels with the factor (q_i-q_j) occurring amongst the $n-1$ factors of (q_i).

On the other hand, when $i=j$, the left-hand side of (22) may be written as

$$\frac{\partial}{\partial q}\prod_{\substack{p=1\\ \neq i}}^{n}(q-q_p),$$

yielding

$$(\mathbf{A}^{-1}\mathbf{A}')_{ii}=\frac{q_i'}{(q_i)}\left[\frac{\partial}{\partial q}\prod_{\substack{p=1\\ \neq i}}^{n}(q-q_p)\right]_{q=q_i}$$

$$=q_i'\frac{\partial}{\partial q}\log\prod_{\substack{p=1\\ \neq i}}^{n}(q-q_p)\Bigg]_{q=q_i}$$

$$=\sum_{\substack{p=1\\ \neq i}}^{n}\frac{q_i'}{q_i-q_p},$$

being the sum of $n-1$ partial fractions.

6. Classification of Transition Points

A *specular transition point* (a reflection or coupling point in physical terms) exists at a point $z=z_0$ in the complex plane at which two or more characteristic roots attain equality.

A transition point of order two exists at a point where $q_1=q_2$; such a point may be denoted by $z=z_{12}$. The first and second diagonal elements each contain one partial fraction with a denominator (q_1-q_2), so this particular partial fraction (but no others down the leading digonal) is singular at the transition point.

Equation (23) shows that every nondiagnonal element is singular at z_{12} along rows one and two only. But only columns one and two have q_1' and q_2' respectively in their numerators; these derivatives may be singular at z_{12}. There is, however, a difference in the order of magnitude of the various coupling terms near the singularity z_{12}. For the coupling coefficients $(\mathbf{A}^{-1}\mathbf{A}')_{ij}$ $(i=1, 2)$ are larger in order of magnitude when $j=1, 2$ than when $j=3, 4, \ldots, n$. This implies that coupling exists mainly only between f_1 and f_2, and that $f_3, f_4, \ldots, f_n$ are relatively free from coupling with f_1 and f_2. In fact, the 2×2 matrix containing the first and second rows and the first and second columns is the distinct factor in determining the coupled solution near z_{12}.

If q_1 and q_2 are expressed in the form

$$q_1=\alpha+\beta,\ q_2=\alpha-\beta,$$

clearly $\beta=0$ at the transition point, while the first part $\alpha=\frac{1}{2}(q_1+q_2)$ is a more irrelevant term in determining the character of the coupling at the transition point.

Similarly, transition points of order two exist whenever $q_a=q_b$; such points may be represented by $z=z_{ab}$. Altogether, $\frac{1}{2}n(n-1)$ sets of second order transition points exist in the z-plane. The solution for f_a and f_b near z_{ab} is locally associated with a 2×2 matrix taken from the a- and bth rows and from the a- and bth columns of $\mathbf{A}^{-1}\mathbf{A}'$.

The determination just considered of the important terms governing the coupling process is clearly unsatisfactory, since infinite terms occur throughout. The embedding transformation discussed in the next sections removes this uncertainty and yields a clearer picture of the coupling process than the matrix $\mathbf{A}^{-1}\mathbf{A}'$, which merely shows where the coupling occurs.

Every nth order equation must of necessity possess these specular transition points of order two, since they are merely given by the solutions of the well-defined equations $q_a=q_b$. But some equations may possess transition points of higher order, not as a matter of necessity but exceptionally. If a point exists at which $q_a=q_b=q_c$, the point z_{abc} is called a third order transition point. All elements in rows a, b, c of the matrix $\mathbf{A}^{-1}\mathbf{A}'$ are now singular at the point z_{abc}, but only the columns a, b, c of these rows prove to be of importance, so the 3×3 matrix taken from these three rows and columns is the determining factor for the coupling beteen f_a, f_b, f_c.

Similar arguments may be used to define an rth $(r\leq n)$ order transition point. Such a point exists if $q_a=q_b=q_c=\ldots$ at $z=z_{abc}\ldots$, where r of the characteristic roots attain equality at this point. The essential coupling between $f_a, f_b, f_c, \ldots$ is determined by the $r\times r$ matrix selected from the appropriate rows and columns of $\mathbf{A}^{-1}\mathbf{A}'$.

These r values of the roots may be expressed according to the scheme (7), but for the purpose of this paper it is assumed that $q_j=\alpha+k^{j-1}\beta$. For the usual case when $r=2$, no assumption is thereby made.

It should be noted that of the remaining $n-r$ roots, s of these may also attain equality at $z=z_{abc}\ldots$. If their common value γ is not equal to α at $z_{abc}\ldots$, then this set of s variables is not coupled to the previous set of r variables.

If various characteristic roots attain equality only as $z\to\infty$ for a certain range of arg z, the coupling is described as *continuous coupling* rather than specular.

Finally, irregular coupling may occur at points where the matrix $\mathbf{P}$ becomes singular, or at points where one or more of the characteristic roots become singular. It is not intended to discuss this form of coupling here, since it does not appear to be amenable to systematic treatment.

7. Embedding Transformation

Specular embedding is defined with respect to a particular transition point z_0 of order $r\leq n$. The first order coupled eqs (18) in the n variables are to be transformed in such a way that the r variables associated with the transition point appear in the new equations with a principal embedded coupled matrix (similar to (10)), while the remaining $n-r$ variables not associated with the transition point appear with a principal uncoupled matrix (a diagonal matrix), in such a way that all coupling terms in all equations are nonsingular at z_0.

The independent propagation of waves not affected by the coupling is thereby exhibited, together with the explicit form of the rth order equation satisfied by the r coupled waves. This coupling is seen to be independent of the remaining $n-r$ waves.

The general coupled equations

$$\mathbf{f}'=\mathbf{Qf}-\mathbf{A}^{-1}\mathbf{A}'\mathbf{f}-\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{Af}$$

are assumed to possess a transition point z_0 at which $q_1=q_2=\ \ldots\ =q_r$. If $q_j=\alpha+k^{j-1}\ \beta$, where $k=e^{2\pi i/r}$, α and β being functions of z such that $\beta=0$ at z_0, a new column $\mathbf{g}$ is defined such that

$$\mathbf{f}=\begin{pmatrix}\mathbf{I}_r\exp\left(\int\alpha dz\right) & 0\\ 0 & \mathbf{I}_{n-r}\end{pmatrix}\mathbf{g}=\mathbf{Eg}$$

say, where $\mathbf{I}_p$ denotes the $p\times p$ unit matrix. Then

$$\mathbf{E}'\mathbf{g}+\mathbf{Eg}'=\mathbf{QEg}-\mathbf{A}^{-1}\mathbf{A}'\mathbf{Eg}-\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{AEg},$$

or

$$\mathbf{g}'=\mathbf{E}^{-1}\mathbf{QEg}-\mathbf{E}^{-1}\mathbf{E}'\mathbf{g}-\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{A}'\mathbf{Eg}-\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{AEg}. \tag{24}$$

Now

$$\mathbf{E}^{-1}\mathbf{QE}-\mathbf{E}^{-1}\mathbf{E}'=\mathbf{Q}-\begin{pmatrix}\alpha\mathbf{I}_r & 0\\ 0 & \mathbf{I}_{n-r}\end{pmatrix}$$

$$=\begin{pmatrix}\beta & 0 & 0 & \ldots & 0 & 0 & \ldots & 0\\ 0 & k\beta & 0 & \ldots & 0 & 0 & \ldots & 0\\ 0 & 0 & k^2\beta & \ldots & 0 & 0 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & 0 & \ldots & k^{r-1}\beta & 0 & \ldots & 0\\ 0 & 0 & 0 & \ldots & 0 & q_{r+1} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & q_n\end{pmatrix}$$

$$=\begin{pmatrix}\beta\mathbf{K}_r & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}$$

say, where $\mathbf{K}_r$ is the diagonal matrix with elements $1, k, k^2, \ldots, k^{r-1}$ and $\mathbf{Q}_{n-r}$ is the diagonal matrix with elements $q_{r+1}, \ldots, q_n$.

Equation (24) now becomes

$$\mathbf{g}'=\begin{pmatrix}\beta\mathbf{K}_r & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}\mathbf{g}-(\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{A}'\mathbf{E}+\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}\mathbf{E})\mathbf{g}. \tag{25}$$

Finally, a new column $\mathbf{h}$ is defined, satisfying

$$\mathbf{g}=\mathbf{R}^{-1}\mathbf{h},$$

where

$$\mathbf{R}=\begin{pmatrix}1 & 1 & 1 & \ldots & 1 & 0\\ \beta & k\beta & k^2\beta & \ldots & k^{r-1}\beta & 0\\ \beta^2 & k^2\beta^2 & k^4\beta^2 & \ldots & k^{2(r-1)}\beta^2 & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ \beta^{r-1} & k^{r-1}\beta^{r-1} & k^{2(r-1)}\beta^{r-1} & \ldots & k^{(r-1)^2}\beta^{r-1} & 0\\ 0 & 0 & 0 & \ldots & 0 & \mathbf{I}_{n-r}\end{pmatrix}$$

$$=\begin{pmatrix}\mathbf{B} & 0\\ 0 & \mathbf{I}_{n-r}\end{pmatrix} \tag{26}$$

say, where $\mathbf{B}$ is an alternant matrix similar to (11).

Equation (25) now becomes

$$\mathbf{R}^{-1\prime}\mathbf{h}+\mathbf{R}^{-1}\mathbf{h}'=\begin{pmatrix}\beta\mathbf{K}_r & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}\mathbf{R}^{-1}\mathbf{h}-(\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{A}'\mathbf{E}+\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}\mathbf{E})\mathbf{R}^{-1}\mathbf{h},$$

or

$$\mathbf{h}'=\mathbf{R}\begin{pmatrix}\beta\mathbf{K}_r & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}\mathbf{R}^{-1}\mathbf{h}-\mathbf{R}\mathbf{R}^{-1\prime}-\mathbf{R}(\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{A}'\mathbf{E}+\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}\mathbf{E})\mathbf{R}^{-1}\mathbf{h}. \tag{27}$$

The original eq (15) is transformed into this form by the transformation

$$\begin{aligned}\mathbf{e}&=\mathbf{S}\mathbf{f}\\ &=\mathbf{P}\mathbf{A}\mathbf{f}\\ &=\mathbf{P}\mathbf{A}\mathbf{E}\mathbf{g}\\ &=\mathbf{P}\mathbf{A}\mathbf{E}\mathbf{R}^{-1}\mathbf{h}\end{aligned} \tag{28}$$

In the next sections, it will be shown that (27) possesses the required embedding properties.

8. Embedded Equation

The coupling matrix occurring in (27) will be considered in two stages, firstly

$$\mathbf{R}\mathbf{R}^{-1\prime}+\mathbf{R}\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{A}'\mathbf{E}\mathbf{R}^{-1} \tag{29}$$

and secondly

$$\mathbf{R}\mathbf{E}^{-1}\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}\mathbf{E}\mathbf{R}^{-1}. \tag{30}$$

Since $\mathbf{R}\,\mathbf{R}^{-1}=\mathbf{I}$, it follows that

$$\mathbf{R}\mathbf{R}^{-1\prime}=-\mathbf{R}'\mathbf{R}^{-1}.$$

Moreover, $\mathbf{E}$ and $\mathbf{R}$ commute owing to the special forms of $\mathbf{E}$ and $\mathbf{R}$; hence (29) becomes

$$
\begin{aligned}
&-\mathbf{R}'\mathbf{R}^{-1}+\mathbf{E}^{-1}\mathbf{R}\mathbf{A}^{-1}\mathbf{A}'\mathbf{R}^{-1}\mathbf{E}\\
&=-\mathbf{E}^{-1}\mathbf{R}'\mathbf{A}^{-1}\mathbf{A}\mathbf{R}^{-1}\mathbf{E}-\mathbf{E}^{-1}\mathbf{R}\mathbf{A}^{-1\prime}\mathbf{A}\mathbf{R}^{-1}\mathbf{E}\\
&=-\mathbf{E}^{-1}(\mathbf{R}'\mathbf{A}^{-1}+\mathbf{R}\mathbf{A}^{-1\prime})\mathbf{A}\mathbf{R}^{-1}\mathbf{E}\\
&=-\mathbf{E}^{-1}(\mathbf{R}\mathbf{A}^{-1})'(\mathbf{A}\mathbf{R}^{-1})\mathbf{E}\\
&=\mathbf{E}^{-1}(\mathbf{A}\mathbf{R}^{-1})^{-1}(\mathbf{A}\mathbf{R}^{-1})'\mathbf{E}. \qquad (31)
\end{aligned}
$$

Also matrix (30) takes the form

$$
(\mathbf{A}\mathbf{E}\mathbf{R}^{-1})^{-1}\mathbf{P}^{-1}\mathbf{P}'(\mathbf{A}\mathbf{E}\mathbf{R}^{-1}). \qquad (32)
$$

At all points where coupling may be neglected (and this includes z_0 since the next section will demonstrate that all coupling matrices are nonsingular at z_0), equation (27) becomes

$$
\begin{aligned}
\mathbf{h}'&=\mathbf{R}\begin{pmatrix}\beta\mathbf{K}_r & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}\mathbf{R}^{-1}\mathbf{h}\\
&=\begin{pmatrix}\mathbf{B} & 0\\ 0 & \mathbf{I}_{n-r}\end{pmatrix}\begin{pmatrix}\beta\mathbf{K}_r & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}\begin{pmatrix}\mathbf{B}^{-1} & 0\\ 0 & \mathbf{I}_{n-r}\end{pmatrix}\mathbf{h}\\
&=\begin{pmatrix}\beta\mathbf{B}\,\mathbf{K}_r\mathbf{B}^{-1} & 0\\ 0 & \mathbf{Q}_{n-r}\end{pmatrix}\mathbf{h}.
\end{aligned}
$$

The elements of $\mathbf{h}$ are thus separated into two distinct parts. If the first r elements are denoted by $\mathbf{h}_r$ and the remaining $n-r$ elements by $\mathbf{h}_{n-r}$, it follows that

$$
\mathbf{h}'_r = \beta\mathbf{B}\,\mathbf{K}_r\mathbf{B}^{-1}\mathbf{h}_r, \qquad (33)
$$

$$
\mathbf{h}'_{n-r}=\mathbf{Q}_{n-r}\mathbf{h}_{n-r}. \qquad (34)
$$

It may immediately be verified that

$$
\begin{aligned}
\beta\mathbf{B}\,\mathbf{K}_r\mathbf{B}^{-1}&=\frac{1}{r}\begin{pmatrix}1 & 1 & 1 & \dots\\ \beta & k\beta & k^2\beta & \dots\\ \beta^2 & k^2\beta^2 & k^4\beta^2 & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\begin{pmatrix}\beta & 0 & 0 & \dots\\ 0 & k\beta & 0 & \dots\\ 0 & 0 & k^2\beta & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\begin{pmatrix}1 & \beta^{-1} & \beta^{-2} & \dots\\ 1 & k^{-1}\beta^{-1} & k^{-2}\beta^{-2} & \dots\\ 1 & k^{-2}\beta^{-1} & k^{-4}\beta^{-2} & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\\
&=\begin{pmatrix}0 & 1 & 0 & \dots\\ 0 & 0 & 1 & \dots\\ \dots & \dots & \dots & \dots\\ \beta^r & 0 & 0 & \dots\end{pmatrix}. \qquad (35)
\end{aligned}
$$

Hence the elements of $\mathbf{h}_r$ are associated with an equation containing a principal embedded coupled matrix, while the elements of $\mathbf{h}_{n-r}$ are associated with a principal uncoupled matrix. In fact,

$$
h_1^{(r)}=\beta^r h_1=vh_1, \qquad (36)
$$

where v vanishes at z_0. The separation of the principal terms has therefore been achieved.

9. Nonsingular Coupling Terms

Every element in the coupling matrices (31) and (32) must now be shown to be nonsingular at z_0.

The elements of $\mathbf{A}\,\mathbf{R}^{-1}$ occurring in (31) are considered. Since

$$\mathbf{AR}^{-1}=\frac{1}{r}\begin{pmatrix}1 & 1 & 1 & \dots\\ q_1 & q_2 & q_3 & \dots\\ q_1^2 & q_2^2 & q_3^2 & \dots\\ \dots & \dots & \dots & \dots\end{pmatrix}\begin{pmatrix}1 & \beta^{-1} & \dots & \beta^{-r+1} & 0 & \dots & 0\\ 1 & k^{-1}\beta^{-1} & \dots & k^{-r+1}\beta^{-r+1} & 0 & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ 1 & k^{-r+1}\beta^{-1} & \dots & k^{-(r-1)^2}\beta^{-r+1} & 0 & \dots & 0\\ 0 & 0 & \dots & 0 & r & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & 0 & 0 & \dots & r\end{pmatrix}$$

it follows that the element $(\mathbf{AR}^{-1})_{ij}$ when $1\leq i\leq n$, $1\leq j\leq r$, is given by

$$r(\mathbf{AR}^{-1})_{ij}=(q_1^{i-1}q_2^{i-1}\dots q_n^{i-1})\begin{pmatrix}1\\ (k^{-1})^{j-1}\\ (k^{-2})^{j-1}\\ \dots\\ (k^{-r+1})^{j-1}\end{pmatrix}\beta^{-j+1}$$

$$=\sum_{p=1}^{r}q_p^{i-1}(k^{-p+1})^{j-1}\beta^{-j+1}$$

$$=\sum_{p=1}^{r}(\alpha+k^{p-1}\beta)^{i-1}k^{-(p-1)(j-1)}\beta^{-j+1}$$

$$=\sum_{p=1}^{r}\left\{\sum_{l=0}^{i-1}{}_{i-1}C_l\alpha^{i-1-l}(k^{p-1}\beta)^l k^{-(p-1)(j-1)}\beta^{-j+1}\right.$$

$$=\sum_{l=0}^{i-1}\left[\sum_{p=1}^{r}k^{(p-1)(l-j+1)}\right]{}_{i-1}C_l\alpha^{i-1-l}\beta^{l-j+1}.$$

Now if $(l-j+1)$ is a multiple of r,

$$\sum_{p=1}^{r}k^{(p-1)(l-j+1)}=r,$$

while if $(l-j+1)$ is not a multiple of r, this sum vanishes by (8). Hence the only terms in β arising in $r(\mathbf{AR}^{-1})_{ij}$ are those of the form β^{Nr} where N is an integer. Moreover, the lowest value of $(l-j+1)$ is $-r+1$, so N can only equal 0, 1, 2, ... Hence

$$(\mathbf{AR}^{-1})_{ij}=\sum_l{}_{i-1}C_l\alpha^{i-1-l}\beta^{Nr}$$

where $l=Nr+j-1$, where N is a non-negative integer chosen so that

$$j-1\leq l\leq i-1.$$

In other words, this element is a polynomial in β^r.

It follows that $(\mathbf{AR}^{-1})'_{ij}$ contains β only through terms such as $Nr\,\beta^{Nr-1}\beta'$.

Finally, since $(\mathbf{AR}^{-1})^{-1}=(\text{adj }\mathbf{AR}^{-1})/(\det\mathbf{AR}^{-1})$ the elements of adj $\mathbf{AR}^{-1}$ are further polynomials in β^r, while

$$\det \mathbf{AR}^{-1} = (\det \mathbf{A})/(\det \mathbf{R})$$

$$= \prod_{\substack{i,j=1 \\ i>j}}^{n} (q_i - q_j)/\beta^{1+2+\cdots+(r-1)} \det \mathbf{K}$$

from (4) and (26). Now when i and j are both less than or equal to r, the relevant factors in the numerator are

$$(q_r - q_{r-1}) \cdots (q_r - q_1)(q_{r-1} - q_{r-2}) \cdots (q_{r-1} - q_1) \cdots (q_2 - q_1)$$

which is proportional to

$$\beta^{r-1}\beta^{r-2} \cdots \beta^2\beta.$$

Hence det $\mathbf{AR}^{-1}$ does not contain β as a factor, so the matrix $\mathbf{AR}^{-1}$ is nonsingular at z_0.

In the product $(\mathbf{AR}^{-1})^{-1}\ (\mathbf{AR}^{-1})'$, each element contains terms of the form $\beta^{Mr}(\beta^{Nr})'$, where $M \geq 0$, $N \geq 0$ are integers. This is proportional to $\beta^{(M+N)r-1}\beta'$, with $M \geq 0$, $N \geq 1$, the lowest power of β yielding $\beta^{r-1}\beta'$. If β, when expanded in terms of $z - z_0$ commences with a term $(z-z_0)^b$, then $\beta^{r-1}\beta'$ contains the factor $(z-z_0)^{br-1}$. This is nonsingular at z_0 if $br - 1 \geq 0$; that is, if $b \geq 1/r$. This implies that $v(z) \equiv \beta^r$, when expanded as a power series in terms of $z - z_0$, should commence with a term that is linear at least.

Hence the elements of the primary coupling matrix (31) are nonsingular at z_0 under these circumstances. The secondary coupling matrix (32) is also nonsingular at z_0, since it has been shown that $\mathbf{AR}^{-1}$ is nonsingular there. It is assumed of course that $\mathbf{P}$ is nonsingular at z_0 in the definition of specular embedding.

10. An Explicit Case

To deal with an explicit case, consider a second order transition point associated with the third order differential equation

$$u''' = v_2 u'' + v_1 u' + v_0 u.$$

Equation (3) is

$$\frac{d}{dz}\begin{pmatrix} u \\ u' \\ u'' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ v_0 & v_1 & v_2 \end{pmatrix}\begin{pmatrix} u \\ u' \\ u'' \end{pmatrix}.$$

The transformation

$$\mathbf{u} = \begin{pmatrix} 1 & 1 & 1 \\ q_1 & q_2 & q_3 \\ q_1^2 & q_2^2 & q_3^2 \end{pmatrix}\mathbf{f}$$

yields equation (5), which, when expanded by the results of section 5, becomes

$$\mathbf{f}' = \begin{pmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{pmatrix}\mathbf{f} - \begin{pmatrix} \dfrac{q_1'}{q_1-q_3}+\dfrac{q_1'}{q_1-q_2} & \dfrac{q_2'}{q_1-q_2}+\dfrac{q_2'}{q_3-q_1} & \dfrac{q_3'}{q_1-q_3}+\dfrac{q_3'}{q_2-q_1} \\ \dfrac{q_1'}{q_2-q_1}+\dfrac{q_1'}{q_3-q_2} & \dfrac{q_2'}{q_2-q_1}+\dfrac{q_2'}{q_2-q_3} & \dfrac{q_3'}{q_2-q_3}+\dfrac{q_3'}{q_1-q_2} \\ \dfrac{q_1'}{q_3-q_1}+\dfrac{q_1'}{q_2-q_3} & \dfrac{q_2'}{q_3-q_2}+\dfrac{q_2'}{q_1-q_3} & \dfrac{q_3'}{q_3-q_1}+\dfrac{q_3'}{q_3-q_2} \end{pmatrix}\mathbf{f}.$$

If z_0 is a point at which $q_1=q_2$, and if $q_1=\alpha+\beta$, $q_2=\alpha-\beta$, the coupling matrix has the form

$$\begin{pmatrix} \frac{\alpha'+\beta'}{q_1-q_3}+\frac{\alpha'+\beta'}{2\beta} & \frac{\alpha'-\beta'}{2\beta}+\frac{\alpha'-\beta'}{q_3-q_1} & \frac{q_3'}{q_1-q_3}-\frac{q_3'}{2\beta} \\ -\frac{\alpha'+\beta'}{2\beta}+\frac{\alpha'+\beta'}{q_3-q_2} & -\frac{\alpha'-\beta'}{2\beta}+\frac{\alpha'-\beta'}{q_2-q_3} & \frac{q_3'}{q_2-q_3}+\frac{q_3'}{2\beta} \\ \frac{\alpha'+\beta'}{q_3-q_1}+\frac{\alpha'+\beta'}{q_2-q_3} & \frac{\alpha'-\beta'}{q_3-q_2}+\frac{\alpha'-\beta'}{q_1-q_3} & \frac{q_3'}{q_3-q_1}+\frac{q_3'}{q_3-q_2} \end{pmatrix}.$$

Singularities arise on account of the terms β', $1/\beta$, and β'/β. If β^2 is approximately linear near z_0, of the form $z-z_0$, $\beta\doteq(z-z)^{\frac{1}{2}}$, so the three singular terms are of the forms $(z-z_0)^{-\frac{1}{2}}$, $(z-z_0)^{-\frac{1}{2}}$ and $(z-z_0)^{-1}$ respectively. The β'/β terms are obviously of a higher order of magnitude than the others; these give rise to the particular matrix

$$\frac{\beta'}{2\beta}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

but other singular elements arise in rows 1 and 2 and columns 1 and 2.

The embedding transformation (28), in which $\mathbf{P}\equiv\mathbf{I}$, becomes

$$\mathbf{u}=\mathbf{AER}^{-1}\mathbf{h}$$

$$=\frac{1}{2}\begin{pmatrix} 1 & 1 & 1 \\ q_1 & q_2 & q_3 \\ q_1^2 & q_2^2 & q_3^2 \end{pmatrix}\begin{pmatrix} \exp\left(\int\alpha dz\right) & 0 & 0 \\ 0 & \exp\left(\int\alpha dz\right) & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & \beta^{-1} & 0 \\ 1 & -\beta^{-1} & 0 \\ 0 & 0 & 2 \end{pmatrix}\mathbf{h}$$

$$=\begin{pmatrix} \exp\left(\int\alpha dz\right) & 0 & 1 \\ \alpha\exp\left(\int\alpha dz\right) & \exp\left(\int\alpha dz\right) & q_3 \\ (\alpha^2+\beta^2)\exp\left(\int\alpha dz\right) & 2\alpha\exp\left(\int\alpha dz\right) & q_3^2 \end{pmatrix}\mathbf{h},$$

while the coupled eqs (27) become

$$\mathbf{h}'=\begin{pmatrix} 0 & 1 & 0 \\ \beta^2 & 0 & 0 \\ 0 & 0 & q_3 \end{pmatrix}\mathbf{h}-\mathbf{E}^{-1}(\mathbf{AR}^{-1})^{-1}(\mathbf{AR}^{-1})'\mathbf{Eh}$$

using results (31) and (35). The principal matrices yield approximately

$$\frac{d}{dz}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ \beta^2 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix},$$

$$h_3'=q_3h_3.$$

exhibiting the embedding property. It should be pointed out that when approximate solutions are written down, it may be necessary to retain some of the coupling terms. For example, in the last equation, an additional term in h_3 must be retained on the right hand side if the W.K.B. approximation for h_3 is to be obtained.

11. A Special Property of the Coupling Matrices

Under the original transformation (16), the coupling matrix is

$$\mathbf{A}^{-1}\mathbf{A}'+\mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A}, \tag{37}$$

while under the embedding transformation (28), the coupling matrix is

$$\mathbf{E}^{-1}(\mathbf{A}\mathbf{R}^{-1})^{-1}(\mathbf{A}\mathbf{R}^{-1})\mathbf{E}+(\mathbf{A}\mathbf{E}\mathbf{R}^{-1})^{-1}\mathbf{P}^{-1}\mathbf{P}'(\mathbf{A}\mathbf{E}\mathbf{R}^{-1}). \tag{38}$$

It is supposed that the last coupling matrix occurs when the first r elements of $\mathbf{f}$ or $\mathbf{h}$ are coupled together at z_0.

However, the $n-r\times n-r$ square matrix formed by crossing out the first r rows and the first r columns of both (37) and (38) are identical. This may be simply demonstrated by partitioning all the matrices between rows r and $r+1$ and between columns r and $r+1$, and expanding the products.

Hence the $n-r$ equations $(r+1\leq j\leq n)$ in the elements of $\mathbf{f}$

$$f_j'=q_jf_j+\text{coupling terms in } f_1, \ldots, f_r+\text{coupling terms in } f_{r+1}, \ldots f_n$$

and the $n-r$ equations $(r+1\leq j\leq n)$ in the elements of $\mathbf{h}$

$$h_j'=q_jh_j+\text{coupling terms in } h_1, \ldots, h_r+\text{coupling terms in } h_{r+1}, \ldots h_n$$

are identical as far as the coupling coefficients multiplying $f_{r+1}, \ldots, f_n$ and $h_{r+1}, \ldots, h_n$ respectively, are concerned.

Hence any coupling diagonal terms not neglected in the equation for f_j' $(r+1\leq j\leq n)$ are identical with the coupling diagonal terms not neglected in the equation for h_j' $(r+1\leq j\leq n)$.

12. Försterling-Type Coupled Equations

As far as ionospheric radio propagation is concerned, two differential equations each of the second order may be coupled together. The original form given by Försterling [1942] and that considered by Budden and Clemmow [1957] are specially simple, in that the characteristic equation for q reduces to a biquadratic. This special feature disguises the more general formulation that is possible.

The general formulation of this new type of coupled equation is as follows. Let z_0 be a transition point of order r, namely a point at which $q_1=q_2=\ \ldots\ =q_r$. Then the n coupled eqs (27) (with no singularities in the coupling coefficients at z_0) may be re-expressed in the form of $n-r+1$ equations in $n-r+1$ variables, taking the form

$$\begin{aligned} h_r^{(r)}&=\beta^r h_r+\text{coupling terms}\\ \mathbf{h}_{n-r}'&=\mathbf{Q}_{n-r}\mathbf{h}_{n-r}+\text{coupling terms.}\end{aligned} \tag{39}$$

The coupling terms involve only the $n-r+1$ variables h_r and $\mathbf{h}_{n-r}$. This result may be shown merely by eliminating $h_1, h_2, \ldots, h_{r-1}$ from the last $n-r+1$ eqs (27), the equation involving h_r' being differentiated $r-1$ times in the process. The coupling terms now involve derivatives of h_r, but the coefficients themselves are also derivatives, which would be small in a slowly-varying medium.

Similarly, if another transition point of order s $(2\leq s\leq n-r)$ occurs such that s of the q's in $\mathbf{Q}_{n-r}$ become equal, the number of equations may further be reduced.

This process is explicitly exhibited for $n=4$, such that a transition point of order 2 occurs at z_0 at which $q_1=q_2$, and another transition point of order 2 at z_1 at which $q_3=q_4$.

The given equation

$$\mathbf{e}'=\mathbf{T}\mathbf{e}$$

is transformed into

$$\mathbf{h}'=\begin{pmatrix}0 & 1 & 0 & 0\\ \beta^2 & 0 & 0 & 0\\ 0 & 0 & q_3 & 0\\ 0 & 0 & 0 & q_4\end{pmatrix}\mathbf{h}+\text{coupling terms}$$

by the transformation

$$\mathbf{e}=\mathbf{PAER}^{-1}\mathbf{h},$$

where

$$\mathbf{E}=\begin{pmatrix}\exp\left(\int\alpha dz\right) & 0 & 0 & 0\\ 0 & \exp\left(\int\alpha dz\right) & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix},\quad \mathbf{R}=\begin{pmatrix}1 & 1 & 0 & 0\\ \beta & -\beta & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

Similarly, if $q_3=\gamma+\delta$, $q_4=\gamma-\delta$, and if

$$\mathbf{F}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \exp\left(\int\gamma dz\right) & 0\\ 0 & 0 & 0 & \exp\left(\int\gamma dz\right)\end{pmatrix},\quad \mathbf{S}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & \delta & -\delta\end{pmatrix}.$$

then the transformation

$$\mathbf{e}=\mathbf{PAER}^{-1}\mathbf{FS}^{-1}\mathbf{j}$$

(in which $\mathbf{E}$, $\mathbf{R}^{-1}$, $\mathbf{F}$, $\mathbf{S}^{-1}$ are mutually commutative) transforms the equation into

$$\mathbf{j}'=\begin{pmatrix}0 & 1 & 0 & 0\\ \beta^2 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & \delta^2 & 0\end{pmatrix}\mathbf{j}+\text{coupling terms}\qquad(40)$$

where the coupling terms are nonsingular at z_0 and z_1.

Now let

$$\mathbf{m}=\begin{pmatrix}j_1\\ j_3\end{pmatrix},\qquad \mathbf{n}=\begin{pmatrix}j_2\\ j_4\end{pmatrix},$$

and two simultaneous matrix equations are formed involving $\mathbf{m}$ and $\mathbf{n}$; that is, eq (40) is multiplied first by $\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}$ and then by $\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$. This yields

$$\mathbf{m}'=\mathbf{n}+\mathbf{Wm}+\mathbf{Xn},\qquad(41)$$

$$\mathbf{n}'=\begin{pmatrix}\beta^2 & 0\\ 0 & \delta^2\end{pmatrix}\mathbf{m}+\mathbf{Ym}+\mathbf{Zn}$$

$$=\mathbf{\Delta m}+\mathbf{Ym}+\mathbf{Zn},\qquad(42)$$

say, where **W**, **X**, **Y**, **Z** are 2×2 matrices originating by partitioning the 4×4 coupling matrix in (40).

Column **m** is now eliminated by differentiating (42):

$$\begin{aligned}
\mathbf{n}'' &= \boldsymbol{\Delta}\mathbf{m}' + \boldsymbol{\Delta}'\mathbf{m} + \mathbf{Y}'\mathbf{m} + \mathbf{Y}\mathbf{m}' + \mathbf{Z}\mathbf{n}' + \mathbf{Z}'\mathbf{n} \\
&= (\boldsymbol{\Delta}+\mathbf{Y})\mathbf{m}' + (\boldsymbol{\Delta}'+\mathbf{Y}')\mathbf{m} + \mathbf{Z}\mathbf{n}' + \mathbf{Z}'\mathbf{n} \\
&= (\boldsymbol{\Delta}+\mathbf{Y})(\mathbf{n}+\mathbf{W}\mathbf{m}+\mathbf{X}\mathbf{n}) + (\boldsymbol{\Delta}'+\mathbf{Y}')\mathbf{m} + \mathbf{Z}\mathbf{n}' + \mathbf{Z}'\mathbf{n} \qquad \text{(from 41)} \\
&= \boldsymbol{\Delta}\mathbf{n} + (\boldsymbol{\Delta}\mathbf{W}+\mathbf{Y}\mathbf{W}+\boldsymbol{\Delta}'+\mathbf{Y}')\mathbf{m} + \mathbf{Z}\mathbf{n}' + (\boldsymbol{\Delta}\mathbf{X}+\mathbf{Y}+\mathbf{Y}\mathbf{X}+\mathbf{Z}')\mathbf{n} \\
&= \boldsymbol{\Delta}\mathbf{n} + (\boldsymbol{\Delta}\mathbf{W}+\mathbf{Y}\mathbf{W}+\boldsymbol{\Delta}'+\mathbf{Y}')(\boldsymbol{\Delta}+\mathbf{Y})^{-1}(\mathbf{n}'-\mathbf{Z}\mathbf{n}) + \mathbf{Z}\mathbf{n}' + (\boldsymbol{\Delta}\mathbf{X}+\mathbf{Y}+\mathbf{Y}\mathbf{X}+\mathbf{Z}')\mathbf{n}
\end{aligned}$$

from (42). $\boldsymbol{\Delta}\mathbf{n}$ is the only term on the right hand side whose coefficient does not involve a derivative. Hence $\boldsymbol{\Delta}$ is the principal uncoupled matrix, while all the other terms are coupling terms.

The coefficient of $\mathbf{n}'$ is $(\boldsymbol{\Delta}\mathbf{W}+\mathbf{Y}\mathbf{W}+\boldsymbol{\Delta}'+\mathbf{Y}')(\boldsymbol{\Delta}+\mathbf{Y})^{-1}$. A subsidiary change of variable will eliminate n_1' from the first equation and n_2' from the second equation if required.

The above process may obviously be carried out for any number of variables.

13. Continuous Coupling

Let the elements of the original matrix **T** (see eq (15)) be functions of z that tend to constant limits as, say, $z \to -\infty$ along the negative real axis. Then the n characteristic roots also tend to limiting values as $z \to -\infty$.

If a range of real z existed such that all the elements of **T** were constants throughout that range, then all the coupling coefficients would be zero, in which case the n coupled eqs (18) reduce to

$$\mathbf{f}' = \mathbf{Q}\mathbf{f},$$

or

$$f_j' = q_j f_j,$$

implying n-independently propagated waves.

More generally, if a domain of z is chosen throughout which the coupling coefficients are small, the coupled equations reduce to

$$f_j' = q_j f_j - (\mathbf{A}^{-1}\mathbf{A}' + \mathbf{A}^{-1}\mathbf{P}^{-1}\mathbf{P}'\mathbf{A})_{jj} f_j,$$

where all nondiagonal coupling terms have been neglected; such approximate equations yield the W.K.B.-type solutions with arbitrary constants. These approximate solutions represent the n characteristic waves in the medium, and they are independently propagated whenever coupling may be neglected. In particular, only one characteristic wave may be considered, namely f_1, with $f_2 = f_3 = \ldots = f_n = 0$. For this particular solution

$$\mathbf{e} = \mathbf{S}\mathbf{f},$$

so

$$e_1 : e_2 : \ldots : e_n = S_{11} : S_{21} : \ldots : S_{n1},$$

thereby providing in this domain the varying ratios of the values of the n original dependent variables associated with this characteristic wave.

The problem demanding attention concerns what happens to these ratios as $z \to -\infty$.

The coupling coefficients may not be neglected when:

i. coupling is specular; the previous embedding theory yields equations dealing with this case;

ii. coupling is irregular; namely, when $\mathbf{P}^{-1}\mathbf{P}'$ or when one of the q's is singular. This paper is not concerned with this case;

iii. coupling is continuous, namely when some of the q's tend to equality as $z\to-\infty$.

To be definite, let $q_1, q_2, \ldots q_r$ have a common limit as $z\to-\infty$. Then the denominators of many of the coupling coefficients considered in section 5 tend to zero. Thus continuous coupling between $q_1, q_2, \ldots, q_r$ is embedded in the background provided by the complete n equations. Previously, specular coupling has been explicitly exhibited without approximations. Here, however, it proves advantageous to use approximations in the equations.

Moreover, it is assumed that the coupling coefficients do not become large in the range of continuous coupling. Every denominator of the form (q_i-q_j), $1\leq i,j\leq r$, contains a derivative in its numerator, and as $z\to-\infty$ it is assumed that the derivative is such that the whole ratio remains finite at least.

Above the region of continuous coupling, a definite approximate solution is considered, namely f_1, with $f_2=f_3=\ldots f_n=0$. Within the region, $\mathbf{e}=\mathbf{S}\,\mathbf{f}$, with $f_{r+1}=f_{r+2}=\ldots=f_n=0$ throughout. Then

$$e_1:e_2:\ldots:e_n=\sum_{j=1}^{r}S_{1j}f_j+\sum_{j=1}^{r}S_{2j}f_j:\ldots:\sum_{j=1}^{r}S_{nj}f_j$$

$$=\sum_{j=1}^{r}S_{1j}(f_j/f_1):\sum_{j=1}^{r}S_{2j}(f_j/f_1):\ldots:\sum_{j=1}^{r}S_{nj}(f_j/f_1) \tag{43}$$

ratios depending upon the $r-1$ values $f_2/f_1, f_3/f_1, \ldots, f_r/f_1$.

These ratios are denoted by $r_2, r_3, \ldots, r_r$ respectively, and their initial values above the region of continuous coupling are all zero.

The coupled equations

$$\mathbf{f}'=\mathbf{Q}\,\mathbf{f}-\mathbf{S}^{-1}\mathbf{S}'\mathbf{f}$$

simplify to

$$f_i'=q_if_i-\sum_{j=1}^{r}(\mathbf{S}^{-1}\mathbf{S}')_{ij}f_j, \qquad (1\leq i\leq r)$$

in which the diagonal elements in $\mathbf{S}^{-1}\mathbf{S}'$ may oftentimes be neglected since these would be small compared with the q_i.

A new set of $r-1$ equations may now be derived, with $r_2, r_3, \ldots, r_r$ as the dependent variables. The differentiation of r_i $(2\leq i\leq r)$ yields

$$r_i'=\frac{d}{dz}\left(\frac{f_i}{f_1}\right)=\frac{f_i'}{f_1}-\frac{f_i}{f_1}\frac{f_1'}{f_1}$$

$$=\frac{1}{f_1}\left[q_if_i-\sum_{j=1}^{r}(\mathbf{S}^{-1}\mathbf{S}')_{ij}f_j\right]-r_i\frac{1}{f_1}\left[q_1f_1-\sum_{j=1}^{r}(\mathbf{S}^{-1}\mathbf{S}')_{1j}r_j\right]$$

$$=q_ir_i-\sum_{j=1}^{r}(\mathbf{S}^{-1}\mathbf{S}')_{ij}r_j-r_i\left[q_1-\sum_{j=1}^{r}(\mathbf{S}^{-1}\mathbf{S}')_{1j}r_j\right],$$

being $r-1$ simultaneous nonlinear differential equations for $r_2, r_3, \ldots, r_r$. These are the appropriate differential equations governing continuous coupling. Integration starts with zero values, yielding $r_2, r_3, \ldots, r_r$ as $z\to-\infty$, which yield in turn the required ratios (43) of the e's associated with this characteristic wave. In the region of continuous coupling, the original ratios $S_{11}:S_{21}:\ldots:S_{n1}$ are converted into the final values.

14. References

Bremmer, H., Some remarks on the ionospheric double refraction, Philips Research Rep. **4,** 189–205 (1949).

Budden, K. G., The theory of the limiting polarization of radio waves reflected from the ionosphere, Proc. Roy. Soc. A, **215,** 215–233 (1952).

Budden, K. G., and P. C. Clemmow, Coupled form of the differential equations governing radio propagation in the ionosphere. II. Proc. Camb. Phil. Soc. **53,** 669–682 (1957).
Budden, K. G., Radio waves in the ionosphere (Cambridge Press, 1961).
Clemmow, P. C., and J. Heading, Coupled forms of the differential equations governing radio propagation in the ionosphere, Proc. Camb. Phil. Soc. **50,** 319–333 (1954).
Försterling, K., Über die Ausbreitung elektromagnetischer Wellen in einem magnetisierten Medium bei senkrechter Incidenz, Hochfr. Elek. **59,** 110 (1942).
Gibbons, J. J., and R. J. Nertney, A method for obtaining the solutions of ionospherically reflected long waves, J. Geophys. Research **56,** 355–371 (1951).
Gibbons, J. J., and R. J. Nertney, Wave solutions, including coupling, of ionospherically reflected long radio waves for a particular E-region model, J. Geophys. Research **57,** 323–338 (1952).
Heading, J., and R. T. P. Whipple, The oblique reflection of long wireless waves from the ionosphere at places where the earth's magnetic field is regarded as vertical, Phil. Trans. Roy. Soc. A, **244,** 469–503 (1952).
Heading, J., Theoretical ionospheric radio propagation (Thesis, Cambridge, England, 1953).
Heading, J., The reflection of vertically-incident long radio waves from the ionosphere when the earth's magnetic field is oblique, Proc. Roy. Soc. A, **231,** 414–435 (1955).
Heading, J., The Stokes phenomenon and certain nth order differential equations, I; preliminary investigation of the equations, Proc. Camb. Phil. Soc. **53,** 399–418 (1957a).
Heading, J., The Stokes phenomenon and certain nth order differential equations, II; the Stokes phenomenon, Proc. Camb. Phil. Soc. **53,** 419–441 (1957b).
Heading, J., The Stokes phenomenon and certain nth order differential equations, III; matrix applications, Proc. Camb. Phil. Soc. **56,** 329–341 (1960).

(Paper 65D6—163)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Worldwide VLF Standard Frequency and Time Signal Broadcasting

A. D. Watt,[1] R. W. Plush,[1] W. W. Brown, and A. H. Morgan

Contribution From the Boulder Laboratories, National Bureau of Standards, Boulder, Colo.

(Received June 15, 1961)

Recent studies and measurements have shown that the phase stability of the signals in the VLF region is very much higher than in the HF spectrum. This fact, along with its excellent coverage characteristics, has caused considerable interest in employing this medium for the wide distribution of standard frequencies and time reference. Basic limitations in stability of the received signals are discussed, including path phase distortion, carrier-to-noise and envelope delay variations as related to precise synchronization of clocks, and highly accurate frequency calibrations.

Also included is a discussion of the present services of standard frequency and time signal stations throughout the world at HF, LF, and VLF.

1. Introduction

The requirement for better standards for precise measurements of frequency and time which are readily available has constantly increased with advances in various fields of science. The scientific unit of time is determined astronomically, and is determined to an uncertainty of a few parts in 10^9 in the course of a year. Since frequency is related inversely to time interval, this uncertainty, strictly speaking, must be transferred to the specification of absolute frequency. Frequency generators, however, stable to parts in 10^{10} or 10^{11} now exist, and form the basis for relative measurements of frequency which for many purposes satisfy the increased requirements mentioned above.

The history of the development of our present system of time is an extremely fascinating subject and it is interesting to note Newton's definition of time, "Absolute, true and mathematical time, of itself, and by its own nature, flows uniformly on, without regard to anything external," [Mach, 1942]. Most of man's time systems have been based on an attempt to find a reference which "flows uniformly on." The very great influence upon our lives and surroundings caused by the revolution of the earth around the sun and the rotation of the earth on its axis has resulted in various definitions [2] of time in terms of years (orbital rotation), days (axial rotation), and subdivision of these units in hours, minutes, and seconds. Time based on apparent solar days (from noon to noon) although apparently satisfactory as a time reference for many applications was shown by astronomical observations to have appreciable nonuniformity.

Since apparent solar days are variable in length with season, mean solar time was devised and the tabulation of this difference between the apparent value and the mean value is called the equation of time. This difference has a maximum value a little in excess of 16 min. Apparent solar days are variable in length with season for two reasons: partly because of the variation in the angular velocity of the sun along the ecliptic in accordance with Kepler's second law of planetary motion, and partly because the inclination of the ecliptic to the celestial equator introduces a variation in the rate at which the projected coordinate of the apparent sun moves along the celestial equator. The mean sun is a fictitious body which moves along the celestial equator at a uniform rate equal to the mean rate of the apparent sun. Because of the fact that a solar time reference (date and time of day, etc.) will depend upon the observer's position, Universal Time (also known as Greenwich mean time) was established with the mean solar time reference based on the prime meridian at Greenwich, England. The problem of relative position for civil purposes has been solved by dividing the earth into various time zones usually differing by one hour with lines of demarcation chosen to minimize the inconvenience caused by crossing time zones in heavily populated areas.

Three kinds of Universal Times are: UT0, which is uncorrected mean solar time; UT1, is Universal Time corrected for observed polar motion and represents the true angular rotation of the earth about its axis;[3] UT2, is Universal Time corrected for both observed polar motion and seasonal variation in speed of rotation of the earth on an extrapolated basis.

The fundamental unit of time was until very recently the mean solar second which was defined as $\frac{1}{86{,}400}$ of a mean solar day. This fundamental unit

[1] Present address of A. D. Watt and R. W. Plush: DECO Electronics, Inc., Boulder Division, 8401 Baseline Road, Boulder, Colo.

[2] An interesting and useful description of time and time signals is contained in "United States Naval Observatory Circular No. 49," published by the U.S. Naval Observatory, Washington 25, D.C. (March 8, 1954). See also "Astronomical Time," G. M. Clemence, Rev. Mod. Phys. 29, 1 (January 1957) and "The Way Things Are," P. W. Bridgman, pp. 135–141 (Harvard University Press, Cambridge, Mass., 1959).

[3] Because of the increased precision with which the earth's rotation can be measured with respect to a reference star rather than the sun, the period of the earth's rotation with respect to the vernal equinox, called a sidereal day, is used for precise determination of the earth's rotation.

of time based upon the mean solar day appeared to be entirely satisfactory until astronomical observations and advances in the fields of communication, electronic frequency standards, and precision instrumentation showed that in fact the earth was not rotating in a strictly constant manner. Even after the corrections of UT1 and UT2, time based on this reference did not flow on uniformly as our definition of it would require. A more uniform astronomical time based on the yearly motion of the earth about the sun, called Ephemeris Time, has long been known. The present definition of the second is $\frac{1}{31,556,925.9747}$ of the tropical year for January 0, 1900 at 12 hr Ephemeris Time [Markowitz, 1959]. This was adopted by the International Committee of Weights and Measures in 1956 and ratified by the 11th General Conference on Weights and Measures in 1960. Discussion of various astronomical and atomic times has been given by Markowitz [1959].

Although quartz crystal clocks were instrumental in revealing many of the variations in rotation of the earth, it was not until atomic standards were placed in use that a new reference of time became conceivable. The use of molecular and atomic spectral lines as frequency standards [Lyons, 1952] has led to the speculation that perhaps here we have a reference which can be made highly independent of external variable influences and as a result may produce a time base of greatly improved uniformity and convenience. The difficulty of keeping molecular and atomic clocks in operation for long periods of time has up until recently prevented their use as a basic time reference. A determination of the atomic transition frequency of cesium in terms of Ephemeris Time by Markowitz, Hall, Essen, and Parry [1958] indicated a value of 9,192,631,770 ± 20 c/s which is the best value available at present. The deceleration in the rotation of the earth about its axis as determined by the cesium standard is in the order of 1.16×10^{-5} radians/year^2, i.e., 0.16 seconds/year^2 [4], which is in good agreement with values obtained by the moon camera of 0.17 seconds/year^2 for that period (June 1955 to June 1958). It is possible [Bullard, 1955] that Ephemeris Time and Atomic Time may not have the same rates; however, further careful experiments will be required to determine this. The stability of the best long time interval bases, which at present are quartz clocks steered by atomic standards is limited to values in the order of one part in 10^{10}. Frequency comparisons quoted to higher precisions are relative, and may be made in two ways. One is to intercompare two oscillators directly for the short term, and the other is to compare one with the mean of a group of three or more of the same quality and assume that their variations are independent, so obtaining a statistical increase in precision for the group. The resonance of cesium is known in terms of the Ephemeris Second to an accuracy of only about 2 parts in 10^9, the limitation being set by the uncertainty in the determination of the Ephemeris Second. Thus, high quality clocks closely steered by atomic standards will not realize the presently defined unit of time with any greater accuracy than this because of the nature of the definition itself. Presently, maximum available relative frequency stabilities for periods of several hours are in the order of: 1 part in 10^{11} for Crystal standards and 1 part in 10^{12} for the Ammonia Maser.

[4] This means that if an atomic clock and "earth" clock were started together, at the end of a year the "earth" clock would be 0.08 sec behind the atomic clock.

2. Basic Limitations of Precision in the Frequency and Time Standards Via Radio Signals

With the availability of very stable time references and frequency standards, it is obvious that considerable care must be exercised if these standards are to be distributed over large areas of the world without appreciable deterioration. If the terrestrial propagation of radio wave energy was at a constant velocity equal to the velocity of light in a vacuum and with a noise free background, it would be a relatively simple matter to distribute frequency and time signals with essentially no loss in stability and accuracy. Under these conditions, the received frequency would be unaffected and the time signals would have a given correctable constant delay of $T'_d=d/v_0$ where: d is the distance and v_0 is the velocity of light $\simeq3\times10^5$ km/s. In practice the radio energy is transmitted along a path with an effective velocity (v) which varies with the characteristics of and conditions along the transmission path so that the actual delay, T_d, may vary also. It is obvious that the time indicated at a given receiving site is $T_r=T_t+T_d$, where the receiving clock can be made to indicate transmitted time T_t by subtracting T_d from the received time T_r during clock setting. At the high frequencies, 2.5 to 30 Mc/s, presently employed for the distribution of frequency and time signals, multiple propagation paths frequently exist with appreciably different delays which can introduce large apparent errors in the received frequency and time. For any given path it may be possible by the proper choice of frequencies to have essentially one dominant mode of propagation; however, even a single mode may have appreciable variations in transmission time.

2.1. Phase Distortion Limitations

In figure 1 the straight dashed line represents the distortionless transmission in free space where the phase delay in radians could be given as

$$\phi'=\frac{2\pi fd}{v_0} \tag{1}$$

where: f is the frequency in cycles per second, and d is the path length expressed in kilometers.

The actual propagation medium, assuming one mode of multiple path to be dominant, may be considered as an electrical network with phase and amplitude versus frequency characteristics which vary as a function of time. The actual time

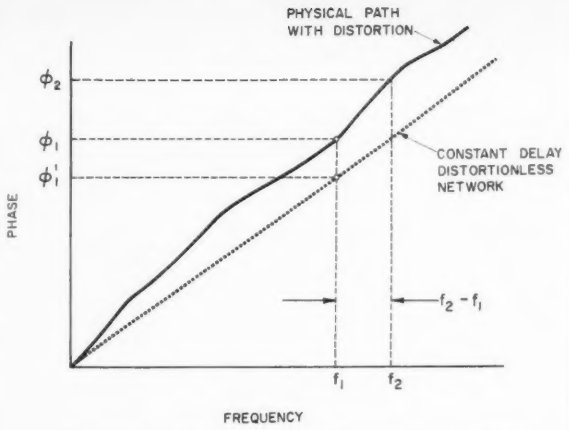

FIGURE 1. *Transmission path phase characteristics.*

delay is

$$T_d = d/v_p \tag{2}$$

where v_p is the average path phase velocity in kilometers per second. It is well known in electrical circuit theory that various types of network delays exist, and if we are to consider the phase delay defined in eq (1), v must be expressed in terms of path phase velocity rather than envelope velocity.

The actual path phase velocity relative to the velocity of light, v_p/v_0, is known to be a function of frequency and the path involved. Important factors may include ionospheric conditions, ground conductivity, and surface roughness. In the VLF region, Jean, Taylor, and Wait [1960] have shown good agreement between experimentally and theoretically determined values of v_p over the 4 to 20 kc/s frequency range. (See also Wait [1961a] and Wait and Spies [1961].) At night they have found values of v_p/v_0 to be $\simeq$1.03 at 4 kc/s, 1.01 at 10 kc/s, and 1.003 at 20 kc/s. Many references pertinent to the velocity of propagation are contained in the paper by Jean, Taylor, and Wait [1960]. In general, v_p increases with surface conductivity, and decreases with increase in ionospheric height and earth's surface roughness.

An actual physical path phase characteristic may look somewhat like the solid wavy curve of figure 1 where a frequency f_1 would be delayed by the phase ϕ_1. This phase delay in seconds can be written as

$$T_d = \frac{\equiv\phi_1}{2\pi f_1} \tag{3}$$

where T_d is the equivalent time delay in seconds. It should be pointed out that the particular phase versus frequency curve illustrated will only apply for a given distance and at a given time and that it may vary in its exact position as a function of time. The effect of these short term variations of phase as a function of time is to limit the precision with which a given standard frequency can be received.

To obtain the value of the received frequency

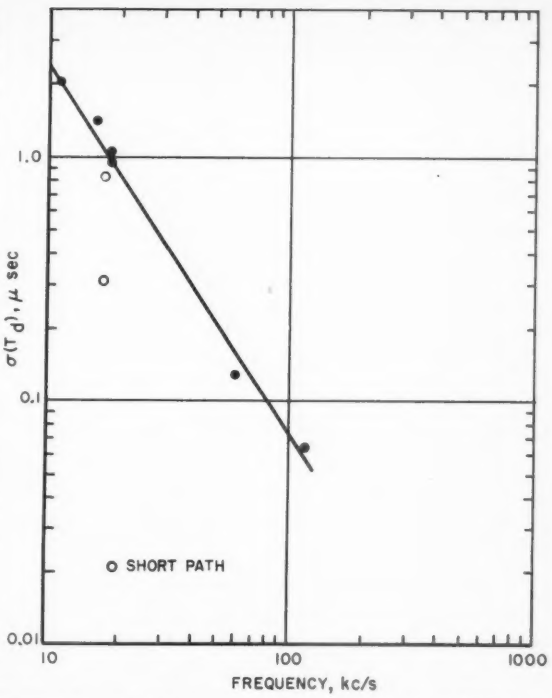

FIGURE 2. *Short term propagation time (phase) stability daytime paths, normalized to single reflection sky wave.*

from the phase measurements it is clear from the definition of frequency that the average departure from the transmitted frequency, $\overline{\Delta f}$, observed in terms of the final and initial phase delays is given by:

$$\overline{\Delta f} = \frac{\phi(T) - \phi(0)}{2\pi T} \tag{4}$$

where T is the time interval, in seconds, over which the phase comparison is made.

The relations governing the transmission path phase stability and signal-integration times required for specific frequency-comparison precisions have been described previously in considerable detail [Watt and Plush, 1959]. In general, where possible, observations should be made near noon at the center of the path. At such time, the phase variations, due to diurnal and random short term effects, are expected to be minimal. Pierce [1957] has experimental evidence over a 5200 km path that the total diurnal phase variation of the 16 kc/s Rugby signal is surprisingly constant throughout the year being in the order of 34 $\pm$1 μsec. Jean[5] over a 7500 km path has observed that the diurnal variation of the phase of the 16 kc/s Rugby signal has substantial variations in pattern. The observed average phase change of $\sim$42 μsec yields about the same effective change in ionospheric height; i.e., $\Delta h \simeq 16$ km using a method described by Wait [1959]. The diurnal patterns observed by Jean

[5] A. G. Jean, private communication. See also C. J. Chilton [1961].

appear to change systematically with seasons and appreciably particularly during ionospherically disturbed conditions. The daytime transmission delay appears to be much less affected by these disturbances than the nighttime phase.

The short time variations in propagation time resulting from path phase instabilities can, in general, be assumed to be randomly distributed about an average time delay $\overline{T_d}$ with a rate of change similar to the fade rate. This means that a single measured value of the delay ranges roughly between $\overline{T_d} \pm \sigma\ (T_d)$, where $\sigma\ (T_d)$ is the standard deviation of the random time delay variable. In general, there will be different values for both $\overline{T_d}$ and $\sigma\ (T_d)$ for day or night paths. Expected standard deviations of short time propagation time delays are shown in figure 2 as a function of frequency for a single reflection daytime skywave in the VLF and LF regions. It should be emphasized that this trend does not continue up into the HF region. The data are obtained from table B–1 [Watt and Plush, 1959] with the additional point at 10.2 kc/s from Tibbals.[6]

2.2. Time Reference and Clock Setting

It is obvious that if a single precisely known uniform frequency were available at a receiving location, that it could be employed in the control of a time standard clock, provided: (a) that the clock could be properly set initially, and (b) that there were no interruptions in the frequency standard or clock. Since neither of these can be guaranteed, it becomes apparent that some method of establishing a time or phase reference at the receiving location is necessary, and that the problem of transmitting a precise time reference is much more difficult than a standard frequency transmission. At the higher frequencies where bandwidth limitations are not severe, this is accomplished by transmitting rather short pulses whose leading edges at the transmitter are quite steep. The delay of these time pulses is also variable but it is not the same as the phase delay described in eqs (2) and (3). Since the pulses are applied in the form of modulation on the rf carrier, the "envelope delay" as defined by Nyquist and Brand [1930] in terms of the slope of the phase characteristic, $d\phi/df$, is the important factor. As we will see in the next section, the HF paths are very instable in both phase and envelope time delay and the timing accuracy obtainable with these pulses is considerably less than that available at the transmitter.

[6] M. L. Tibbals, private communication.

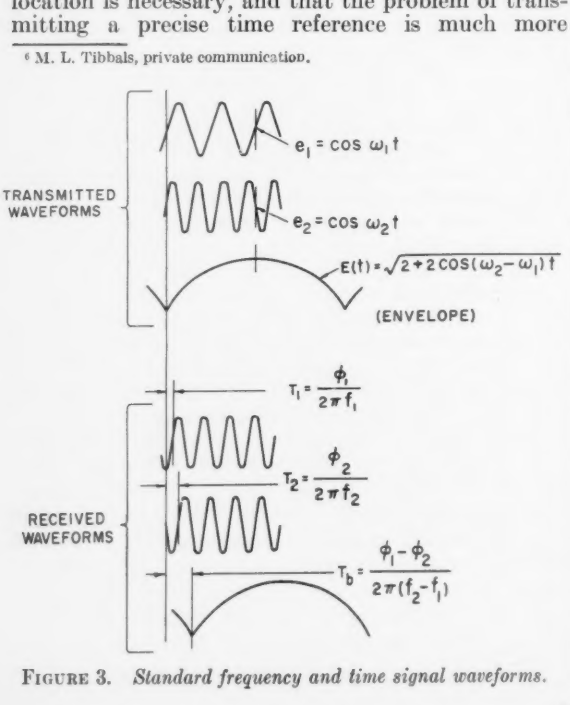

FIGURE 3. *Standard frequency and time signal waveforms.*

In view of the much more stable phase delay conditions in the VLF region as described by Pierce [1957] and Allan, Crombie, and Penton [1956] along with its excellent coverage characteristics, considerable interest has been exhibited in employing this medium for the distribution of standard frequencies and time reference. The channel bandwidths available in this frequency region, along with the very narrow bandwidths of the transmitting antenna, prevent the use of short-time pulses for the establishment of coarse time markers which could be used in identifying a particular rf cycle. An alternate method employing two closely spaced and alternately transmitted related frequencies is being considered. A description of this method is given by Morgan [1961]. The transmitted and received waveforms for such a narrow band time signal system are shown in figure 3 where it is apparent that the envelope of the voltage difference produces time reference markers with ambiguities spaced by the reciprocal of the frequency difference. Coarse time markers may, if desired, be provided by the times of frequency switching. On-off keying can also be considered for date time code. In a dual frequency system, transmitted waveforms are given as:

$$e_1 = \cos \omega_1 t \tag{5}$$

$$e_2 = \cos \omega_2 t \tag{6}$$

while the envelope of the phasor sum is given as

$$E(t)_{\text{envelope}} = \sqrt{2+2\cos(\omega_2-\omega_1)t} \tag{7}$$

It can be observed that at the transmitter, zero time is our phase reference. At the receiver, the phases of e_1 and e_2 are delayed and these two signals are

$$e_{1r} = \cos[\omega_1 t + \phi_1] \tag{8}$$

$$e_{2r} = \cos[\omega_2 t + \phi_2] \tag{9}$$

Since these two signals can be employed to phase lock high quality local oscillators, it is possible to have two highly constant amplitude sinusoids as indicated. The envelope voltage waveform at the receiver is then

$$E(t)_{\text{envelope received}} = \sqrt{2+2\cos[(\omega_2-\omega_1)t+(\phi_2-\phi_1)]} \tag{10}$$

where it is evident that the important phase is now

$\phi_2-\phi_1$ and the resulting time delay is

$$T_b=\frac{\phi_2-\phi_1}{2\pi(f_2-f_1)} \tag{11}$$

where T_b is the envelope delay in seconds.

If the individual frequencies concerned are separated far enough, the path phase fluctuations become independent and the standard deviation of T_b would be

$$\sigma(T_b)=\frac{\sqrt{2}\sigma(\phi)}{2\pi(f_2-f_1)} \tag{12}$$

assuming that $\sigma(\phi_1)=\sigma(\phi_2)=\sigma(\phi)$.

It can be seen that the time variations will decrease with an increase in frequency separation. It should be emphasized that as the frequency separation is reduced the relative phases tend to become correlated and as a result the envelope reference time variation does not increase without limit since the phase difference tends to approach zero. Measured standard deviations of envelope delay variations for a 1 kc/s frequency spacing in the VLF band are shown by Casselman, Heritage, and Tibbals [1959] to be in the order of ± 20 μsec. (See also Stone, Markowitz, and Hall, 1960.)[7]

Employing the ionospheric roughness parameter described by Watt and Plush [1959] the standard deviation of short term phase variation; i.e., $\sigma(\phi)$, of a VLF carrier received over a 4200 km path is found to be in the order of 0.09 radians. When this value is used in (12) the expected $\sigma(T_b)$ expressed in microseconds is shown in figure 4. Experimental values for $\sigma(T_b)$ obtained from Heritage and Tibbals [8] for spacing of 200, 1,000, and 3,000 c/s appear to substantiate the general level and trend of this curve.

In order to permit identification of a specific cycle of the carrier frequency f_1, it is obvious that $\sigma(T_b)$ must be less than $0.5/f_1-\sigma(T_d)$ which at 20 kc/s will be in the order of ± 22 μsec. Until such time as the expected reduction in $\sigma(T_b)$ due to integration for time intervals long compared to the fade periods is obtained, it appears that frequency separations in the order of 1 kc/s will be required for carrier cycle identification in the 20 kc/s region. If the frequency separation chosen is rather large, the (f_2-f_1) markers may be difficult to resolve from the on-off keying envelope. This is due in part to the normal build up time required by the narrow antenna bandwidths, see typical shapes in figure 5, and the variation in position of this envelope as received due to path envelope delay. Should the on-off keying phase markers received with time delay variations $\sigma(T_e)$ not be stable enough to resolve (f_2-f_1) markers, it may be necessary to employ an additional frequency providing an intermediate reference, viz $(f_3-f_1)<(f_2-f_1)$.

Since the precision of frequency comparison is defined as the phase jitter divided by the observing period T (7.3 of Watt and Plush, 1959) we can write

$$\epsilon=\sqrt{2}\sigma(\phi)/2\pi fT. \tag{13}$$

The time delay for a given phase delay is $T_d=\phi/2\pi f$, and since $\sigma(T_d)=\sigma(\phi)/2\pi f$, we can write (13) as

$$\sigma(T_d)=\epsilon T/\sqrt{2}. \tag{14}$$

Figure 6 shows the manner in which this relationship can be employed to determine the variations expected in a given clock relative to a standard clock T seconds after it has been set to coincidence with the standard clock. It is interesting to note that if an all daylight radio path such as the 16 kc/s transatlantic path observed by Pierce [1957] is employed, the maximum standard deviation of time difference is less than 2 μsec.

It should be pointed out that the problem of subtracting out the mean true transmission time, $\overline{T_d}$,

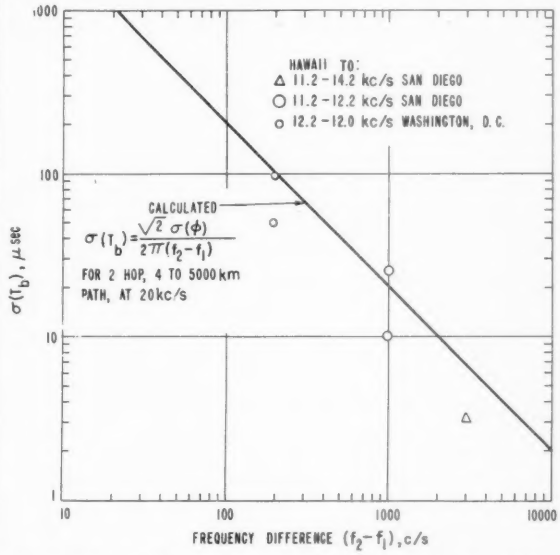

FIGURE 4. *VLF transmission path envelope phase stability.*

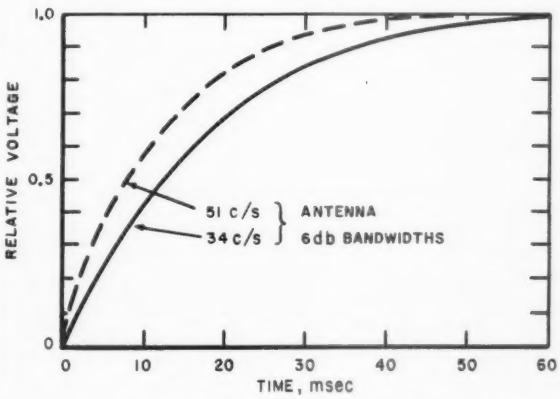

FIGURE 5. *Keying response envelope of typical VLF transmitting antenna.*

[7] Stone, Markowitz, and Hall in a recent paper [1960] have indicated VLF time comparisons of 500 μsec at Washington, D.C., from NBA in Panama. Employing a 6 db transmitting bandwidth of 44c/s on fig. 4 and assuming that the envelope matching would approach that of two frequencies spaced by the 6 db bandwidth we obtain $\sigma(T_b)\simeq 450$ μsec for a path of comparable length. Actually the spacing chosen should likely be less which would yield better agreement.

[8] Private communication.

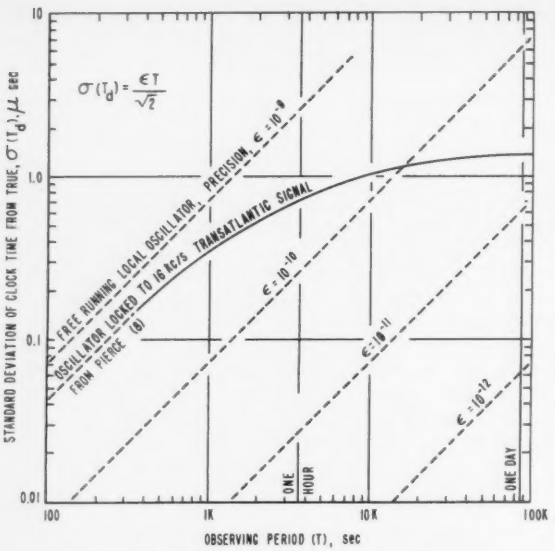

FIGURE 6. *Clock-time deviations from original setting as a function of observing period.* ([8] Pierce 1957).

whether it be a phase delay or envelope delay, still exists. Ionospheric disturbances caused by solar related proton storms or meteoric showers have the largest apparent effect on the nighttime phase delay [Chilton, 1961]. As a result, the midday values of $\overline{T_d}$ are expected to be relatively constant from day to day with the exception that solar flares effect the day values. Over some paths, it is possible that there may be appreciable seasonal variations in even the midday value of $\overline{T_d}$. Additional studies are required before the errors likely in predicting $\overline{T_d}$ can be specified as closely as desired. If two or possibly three VLF standard frequency and time signal stations are operating simultaneously at different rf frequencies but with a common frequency base as well as a common time reference, it is expected that the average delays due to transmission time can be obtained with greater accuracy. Once the average path delays have been subtracted out, a clock at a given receiving site can be set with a high degree of precision (probably limited mainly by the ability to determine absolute average path delay), to the time generated by the network of VLF standard frequency and time signal broadcasting stations.

2.3. Carrier to Noise Limitations

The limitations of noise to precision of frequency comparison have been previously described [Watt and Plush, 1959] and we shall only calculate the effects of rms carrier to rms thermal noise ratio on time signal accuracy.

From equation (C1) of Watt and Plush [1959] and (3)

$$\underset{\text{(noise)}}{\sigma(T_d)} = \frac{N/C}{2\sqrt{2}\pi f_1} \tag{15}$$

where $\sigma(T_d)$(noise) is now the standard deviation of the time reference in seconds derived from the carrier frequency (f_1) phase, and N/C is the rms noise to carrier ratio in the receiver. With post detection filtering or integration following a linear detector, N should be calculated in the effective bandwidth of the integrator. If we consider noise density or in particular, the noise voltage per 1 kc/s bandwidth concept where $N_{1\text{ kc}}$ has the dimensions volts/$\sqrt{\text{kc/s}}$, we obtain

$$\underset{\text{(noise)}}{\sigma(T_d)} = \frac{\sqrt{B_r/1{,}000}}{2\sqrt{2}\pi f_1 C/N_{1\text{ kc}}} \tag{16}$$

where B_r is the effective receiver bandwidth in cycles per second.

When timing is derived by an envelope beat pattern between two carriers, the effects of noise are given by

$$\underset{\text{(noise)}}{\sigma(T_b)} = \frac{N/C}{2\pi(f_2 - f_1)} \tag{17}$$

or

$$\underset{\text{(noise)}}{\sigma(T_b)} = \frac{\sqrt{B_r/1{,}000}}{2\pi(f_2 - f_1)C/N_{1\text{ kc}}} \tag{18}$$

where it is obvious that the time stability for these conditions is poorer for a given $C/N_{1\text{ kc/s}}$ ratio than obtains for $\sigma(T_d)$ (noise).

We can also determine the stability of a time reference obtained from a keyed envelope obtained at the one half amplitude point of the leading edge with coherent detection. For a single sample this is shown in the appendix to be

$$\underset{\text{(noise, sampled)}}{\sigma(T_e)} = \frac{1}{(C/N)2B_t} \tag{19}$$

where: $\sigma(T_e)$ is in seconds and B_t is the transmitter 6 db bandwidth in c/s. If the carrier is keyed on periodically (for example, once each second) the resulting envelope reference can be integrated to reduce the effects of noise. The improvement is equal to the square root of the number of individual pulses, and if the repetition rate is F

$$\underset{\text{(noise, integrated)}}{\sigma(T_e)} = \frac{1}{(C/N)B_t\sqrt{TF}} \tag{20}$$

or

$$\sigma(T_e) = \frac{\sqrt{B_r/1{,}000TF}}{(C/N_{1\text{ kc}})B_t} \tag{21}$$

where: T is the integration period, TF is the number of samples averaged, B_r is the effective receiver bandwidth in c/s. It is apparent from (21) that the time reference derived from the envelope leading edge improves directly as the carrier to noise ratio and as the first power of the transmitting antenna bandwidth and as the square root of the

integration time. In all the derivations involving the rms carrier to noise ratios, the significant noise background considered is of thermal type. The actual background will normally be atmospheric noise which in a wide bandwidth has a very different statistical nature [Watt and Maxwell, 1957]. The effects of atmospheric noise upon radio systems are in general quite different from those of thermal noise [Watt, Coon, Maxwell, and Plush, 1958]. For a pulse sampling system with integration, the required rms C/N may be appreciably less than indicated in eqs. (20) and (21). Hefley[9] [1960] for example has found with Loran-C systems that for equivalent performance the carrier to atmospheric noise ratio can be much less than the carrier to thermal noise ratio; however, it must be remembered that this system employs a very wide RF bandwidth.

When carrier phase or frequency stability is considered, the narrow bandwidths (0.1 c/s or less) employed will make the atmospheric noise statistics the same as thermal noise [Watt and Maxwell, 1957] and in these cases the formulas given (15) through (18) should apply if linear receivers are employed.

3. Present Services

There are at present quite a large number of standard-frequency and time-signal stations in use throughout the world. VLF, LF, and HF are employed for large area coverage with some higher frequencies being used for local distribution. A description of the stations operating along with their location and characteristics is given in Annex I of CCIR Report 66 [1956].

Annex II of the same document describes the characteristics of projected standard frequency and time signal stations at four additional locations. In general, the high frequency stations all employ pulse modulation of the carrier frequency for time signal transmissions and many of them in addition have standard audiofrequencies employed as modulation at various times. The stability as broadcast is usually very good, for example with the National Bureau of Standards station WWV [1960a] the carrier and audiotones have stability of one part in 10^9 at all times with normal daily deviations of less than two parts in 10^{10}. The time signals have essentially the same stability and the maximum deviation from UT–2 is about ± 50 msec. Final corrections to the frequency as broadcast versus the U.S. Frequency Standard are published monthly by the National Bureau of Standards [1960b] in the Proceedings of the IRE while final corrections to the time signals versus UT–2 time are published by the U.S. Naval Observatory.

3.1. HF Standard Frequency and Time Services

Because of the variation in the propagation medium at HF, the frequency as received is generally much less stable than that transmitted. Actually, for high precision measurements it is frequently necessary to use long averaging periods of 10 to 30 days. Ideally, the following conditions should prevail at the time measurements are being made; (a) all daylight (or darkness) over radio path, (b) no ionospheric disturbance in progress, (c) no part of the radio path should pass near either auroral zone, (d) single mode propagation should be existing. In general, frequency determinations can be made either by a direct comparison or by means of the time pulses. Some idea of the precision of frequency comparison obtainable can be seen from figure 7 which shows the precisions available from WWV and WWVH over paths ranging from approximately 2,000 to 8,000 km. Line A shows that for short period samples, such as 100 sec, the precision available is in the order of 2 parts in 10^7. If similar 100-sec samples are made each day, after 30 days the precision has improved to one part in 10^8. It should be emphasized that this type of frequency comparison does not make full use of the transmissions and that the improvement with observing period is not as great as would be predicted if all the actual information received was fully utilized for the total periods shown. Line B shows the results obtained at Boulder employing time pulses compared for approximately 10 sec near noon each day and averaged by means of a clock. Here the daily averaging time is 24 hr. The point shown for 30 days is arrived at as follows. Daily comparisons of the local clock versus the received time of WWV are used to obtain running 10-day averages of time difference. Each day one new daily value is added and the 11-days old value is dropped and the re-

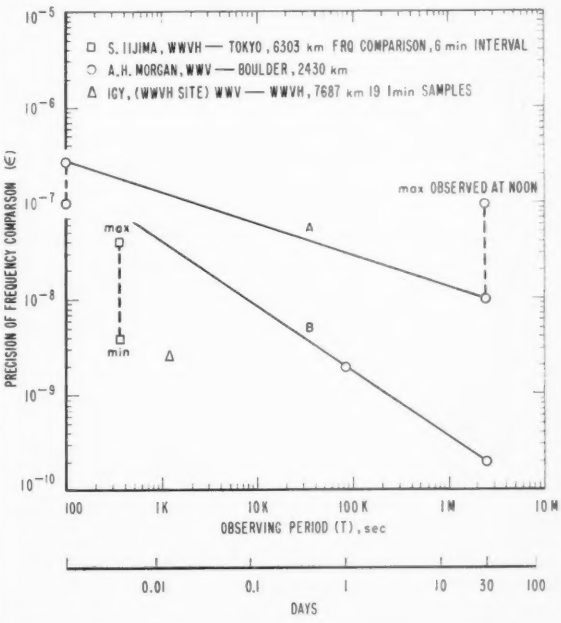
□ S. IIJIMA, WWVH — TOKYO, 6303 km FRQ COMPARISON, 6 min INTERVAL
○ A.H. MORGAN, WWV — BOULDER, 2430 km
△ IGY, (WWVH SITE) WWV — WWVH, 7687 km 19 1min SAMPLES
max OBSERVED AT NOON
max
min
A
B
PRECISION OF FREQUENCY COMPARSON (ε)
OBSERVING PERIOD (T), sec
DAYS

FIGURE 7. *Frequency comparison precision at 10 Mc/s.*

Line A=100 sec samples, frequency comparisons 1 sample/day at noon.
Line B=24-hr average each day, time pulses, compared at 10 a.m. local time.
All values are standard deviation and at best time of day (disturbed day data not included).

[9] Private communication.

sultant 10 time differences averaged. The 30-day running averages of the above 10-day averaged values are obtained in a similar manner. The precisions obtained are appreciably better, being in the order of ± 2 parts in 10^9 for 1 day and approximately ± 2 parts in 10^{10} for 30 days.

3.2. LF Standard Frequency Transmissions

At present there are at least six low frequency standard transmissions being operated including MSF at 60 kc/s in Rugby, England; WWVB at 60 kc/s in Boulder, Colo.; DCF77 at 77.5 kc/s in Mainflingen, Germany; OLP at 50 kc/s in Czechoslovakia; HBJ at 96.04 kc/s in Switzerland; and A5XA at 133⅓ kc/s at Fort Monmouth, N.J., U.S.A. Other characteristics of these transmissions are described in Annex III of CCIR Report 66 [1956]. Pierce [1957] has shown the standard deviation of frequency comparison for MSF transmissions on a transatlantic path to be in the order of 1 times 10^{-10} for a 30-minute observing period. Further calculations for LF transmissions are compared with experimental results by Watt and Plush [1959] and shown here in figure 8. Pierce [1958] has shown average frequency comparison precisions from WWVB (60 kc/s) as received at Cruft Laboratory, of 1.4 parts in 10^{11} in an observing period of about one hr. From equation (14) this would correspond to a short term time variation $\sigma(T_d) \simeq 0.36$ μsec.

Recent plans to distribute time signals via Loran-C navigation system, pulse transmissions at 100 kc/s, have been described by Doherty [1961]. Very precise time signal distribution in the order of 0.2 μsec is possible within the ground wave coverage range of about 1,500 km. By employing sky wave modes, timing precision in the order of 1 μsec appears likely out to about 3,500 km.

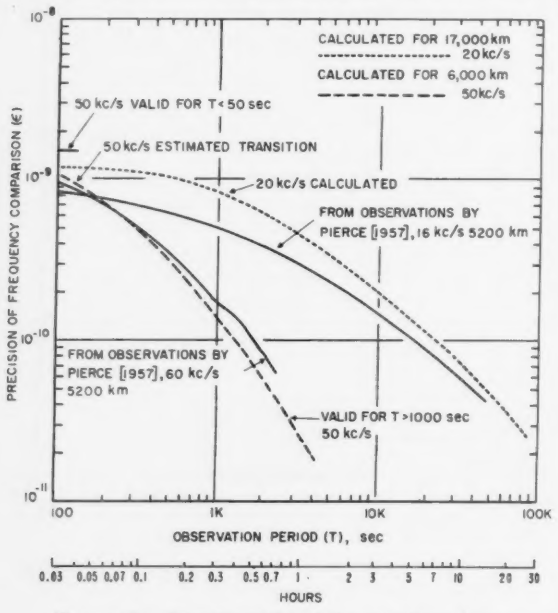

FIGURE 8. *Experimental and theoretical precisions.*

NOTE: Due to diurnal phase change observing periods of longer than 10 K include several days.

It is obvious from this and the preceding sections that in a limited area the LF systems (50 to 100 kc/s) will provide for more precise comparison of frequency and time than the VLF systems (15 to 20 kc/s). The propagation loss, however, is much higher in the 50 to 100 kc/s band which restricts the useful range to distances of about 3,000 to 4,000 km. LF transmissions at 60 kc/s in England [Steele, 1955] have shown that timing can be obtained to ± 15 μsec by pulse envelope methods.

3.3. VLF Standard Frequency Transmissions

Annex III of CCIR Report 66 [1956] describes the transmissions from GBR at 16 kc/s which has been used as a basis for much of the pioneer work on standard frequency broadcasting in the very low frequency region. The precisions available from this transmission, and GBZ on 19.6 kc/s, based primarily on observations by Pierce [1957, 1958] and Allan, Crombie and Penton [1956], are included in sections 2.2 and 4. Stabilized transmissions with phase locked master-slave relationship from Hawaii and San Diego have also served as a valuable source of information for VLF path phase stability as described by Casselman, Heritage, and Tibbals [1959]. Recently interim station broadcasts of the U.S. frequency standard have been initiated from WWVL on 20 kc/s from near Boulder, Colo. Time and stabilized frequency broadcasts from NBA at 18 kc/s also recently have been initiated.[10] Additional VLF transmissions are in the process of being stabilized in frequency which should make them available for path phase stability measurements. In the VLF region both theory and experiment show that greater observing times are required for a given degree of frequency comparison precision than is true in the LF region; however, when long range paths are considered the low attenuation rates obtaining in the VLF band make emissions in this frequency region useful for worldwide coverage.

4. Characteristics and Expected Coverage for a VLF Standard Frequency and Time Signal Broadcasting Station

Based on the previous analysis [Watt and Plush, 1959] the precision of frequency comparison expected as a function of observing period T was prepared and is shown as figure 8 [Pierce, Winkler, and Corke, 1960].[11] It is obvious when comparing figures 7 and 8 that a VLF broadcast will provide one or two orders of magnitude improvement over present HF broadcasts in precision of frequency comparison for a given observing period. This great increase in performance has clearly indicated the desirability of designing and constructing a high quality VLF standard frequency broadcasting station or network of 2 or 3 stations whose primary

[10] Precise time and stabilized frequency broadcasts at 18 kc/s over NBA, Summit, Canal Zone, were announced in U.S. Naval Observatory Time Service, Notice No. 8, 18 November 1959. See also a recent paper by Stone, Markowitz, and Hall [1960].

[11] A recent paper by Pierce, Winkler, and Corke [1960] describes results at 16 kc/s over a 5,200 km path that indicate possible precisions of 2×10^{-11}.

function is to provide extremely precise standard frequency and time signals.

Because of the desirability of having such a station near the U.S. standard of frequency located at Boulder, Colo., a study has been made to determine design parameters and service obtainable at various locations. Seventeen different receiving sites were chosen and the total number of measurements obtainable each day for 99 percent of all hours were calculated based on a radiated power of 100 kw. Calculations were made for a range of radiated powers from 25 to 100 kw and the resulting number of measurements obtainable as a function of the cost per measurement was determined. The results of this study [Watt and Plush, 1959] indicated a preference for a radiated power in the order of 100 kw.

The results of this study are summarized in table 1, where the various assumptions made are indicated along with a tabulation of the total number of measurements obtainable for 99 percent of all hours with 100 kw radiated at 20 kc/s [12] for the months of June and December assuming a precision of frequency comparison of 1 part in 10^9. The method of analysis and some of the assumptions made in this analysis are contained in the following sections.

4.1. Propagation Path Attenuation Rates

It is relatively well known that the attenuation rate (a) to be employed in field strength equations such as (2) in the paper by Watt and Plush [1959], expressed in decibels per 1,000 km (db/K), is dependent upon a number of factors including: frequency, ionospheric conditions, earth's surface conditions, and the earth's magnetic field.

Experimental and theoretical studies indicate that there is a broad minimum in the attenuation-versus-frequency curve which is centered somewhere in the vicinity of 16 to 18 kc/s. The shape of this attenuation curve is likely to vary between day and night conditions, and is also likely to be different for east-to-west and west-to-east propagation. It would appear from an analysis of results obtained from many different sources that the east-to-west (magnetic) attenuation rate in some cases may be as much as 1 or 2 db/K greater than the west-to-east attenuation rates. This effect is expected to be more pronounced at night; in addition this difference in attenuation rate with direction is expected to be greater at the lower frequencies; viz, 10 to 14 kc/s, and relatively small in the vicinity of 20 kc/s. These effects have been discussed by various authors [H. J. Round, T. E. Eckersley, K. Tremellen, and F. C. Lunnon, 1925; K. G. Budden, 1951 and 1952; J. R. Wait, 1958 and 1961b; and D. D. Crombie, 1958].

Approximate values for the average daytime attenuation rate (a) in the vicinity of 20 kc/s are: sea $\simeq 2$, average land $\simeq 4$, estimated values for permafrost and icecap are 7 and 18. When mixed surface conditions are encountered along the path, as is usually the case in practice, the total average path attenuation will not always be exactly equal to the sum of the attenuations expected from each individual portion of the path. This may be due in some cases to a transfer of energy from the dominant to higher order modes caused by the discontinuities in conductivity at the surface. It is also likely that low conductivity at the earth's surface will produce appreciably greater attenuation during the day when the ionosphere is low than at night when it is higher.

After determining the various amounts of surface types along the individual paths, expected average attenuation rates have been assigned for each of the receiving locations in table 1. In some cases reciprocal path values have been employed and, in general, values quoted are as high as anticipated from a consideration of all possible attenuation mechanisms.

Field strengths anticipated for these particular paths have been calculated employing the relationships given by Watt and Plush [1959]; the results are shown in column 5 of table 1 for an assumed radiated power of 100 kw at a frequency of 20 kc/s.

4.2. Expected Availability and Precision of Frequency Comparison

The carrier to noise available at the receiving location may in some cases be the limiting factor as regards the time required to obtain a frequency comparison of, say, 1 part in 10^9. Median atmospheric noise fields expected for the various receiving locations have been obtained from Crichlow, Smith, Morton, and Corliss [1955] and the results presented in columns 6 and 7 of table 1. Combining the previous information regarding carrier field intensities with the noise field intensity data, we have obtained the minimum expected observation times required to compare frequency to 1 part in 10^9 for 100 kilowatts radiated as shown in columns 14 and 15.

Pierce [1957] has shown that the transmission path phase is very stable and that the diurnal variations caused by sunrise or sunset along the Rugby to Cambridge path are surprisingly constant throughout the year. During the all sunlight period the path phase appears more stable than at night for frequencies of 16 to 60 kc/s. We have determined, on the basis described by Watt and Plush [1959], the minimum observing times required to obtain a precision of 1 part in 10^9 based on path phase variations expected for both day and night conditions. It should be noted that $\sigma(\phi)$ has been assumed to increase as the square root of distance which is what one would expect of sections of variable delay wave guides connected in series. These results are presented in columns 12 and 13.

Although it is obvious that in a case where minimum observing times based on path phase variability and carrier-to-noise ratio are nearly the same, the combined effect will require a longer observation time than indicated for either limitation. We have neglected this effect and have considered only the one condition which requires the longer period.

In order to get some idea of the relative service obtainable from the proposed standard frequency and time signal broadcast, the total number of

[12] It should be noted that 20 kc/s has been employed in these calculations because: (1) 20 kc/s is an internationally assigned standard frequency, (2) the power requirements for worldwide coverage are expected to be near a minimum at this frequency, (3) considerable propagation data is available in this region. Further research on attenuation and phase stability should be conducted over the whole VLF spectrum to determine the suitability of other frequencies for standard frequency and time distribution.

TABLE 1 – STANDARD VLF BROADCAST STATION. ANTICIPATED RECEPTION CAPABILITIES – 100 kW AT 20 kc/s RADIATED NEAR BOULDER, COLO.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		17	
No	LOCATION	DIST, km	(ESTIMATED) PATH ATTEN DAYLIGHT db/1000 km	CARRIER FIELD FOR 100 kW RAD db RELATIVE TO 1 μv/m	MEDIAN NOISE FIELD db RELATIVE TO 1 μv/m, 1 kc/s BAND		hr ALL DAYLIGHT ON PATH	No	hr ALL NIGHT ON PATH	No	MIN OBSERVATION TIME – hr FOR ε=10⁻⁹ BASED ON PATH PHASE VARIATION		MIN. OBSERVATION TIME – hr FOR ε=10⁻⁹ BASED ON C/N WITH 100 kW RADIATION		TOTAL No. OF MEASUREMENTS OBTAINABLE FOR 99% OF ALL HOURS, 100 kW			
					JUNE	DECEMBER	JUNE	DECEMBER	JUNE	DECEMBER	DAY	NIGHT	JUNE	DECEMBER	JUNE			DECEMBER	
															DAY	NIGHT	TOTAL	DAY	NIGHT	TOTAL
1	WASHINGTON, D.C.	2 370	5.0	59.1	45.0	35.0	12.8	7.2	7.2	12.8	0.001	0.07	0.0145	0.0069	882	102	984	1 042	182	1 224
2	MONTREAL, Canada	2 370	5.0	59.1	40.0	32.0	12.9	6.7	6.7	12.9	0.001	0.07	0.01	0.0054	1 289	95	1 384	1 239	184	1 423
3	THULE, Greenland	4 470	5.5	43.7	30.0	26.0	14.8	4.2	4.2	14.8	0.002	0.14	0.015	0.011	985	29	1 014	380	104	484
4	KEFLAVIK, Iceland	5 770	5.0	38.7	32.0	30.0	11.48	0/(1)	11.48	0.006	0.18	0.026	0.022	0.0082	440	4	444	44	60	104
5	RUGBY, England	7 400	4.5	33.8	44.0	34.0	8.3	1.5	1.5	8.3	0.01	0.24	0.096	0.044	85	5	90	33	32	65
6	PARIS, France	7 770	4.5	31.7	40.0	35.0	8.0	1.4	1.4	8.0	0.01	0.28	0.084	0.057	94	4	98	23	27	50
7	TORINO, Italy	9 000	4.5	26.0	42.0	35.0	7.1	1.1	1.1	7.1	0.02	0.3	0.15	0.088	46	2	48	111	22	33
8	MOSCOW, U.S.S.R.	8 880	5.5	30.5	40.0	35.0	6.3	-1.7	-1.7	6.3	0.02	0.3	0.09	0.062	69	–	69	–	20	20
9	NEW DELHI, India	12 850	5.0	-0.1	45.0	40.0	2.6	-2.0	-2.0	2.6	0.5	1.0	1.4	0.96	1	–	1	–	2	2
10	FAIRBANKS, Alaska	3 880	5.0	39.5	30.0	28.0	14.3	4.1	4.1	14.3	0.002	0.12	0.0098	0.0082	1 458	32	1 490	499	116	615
11	SAN DIEGO, California	1 450	5.0	65.6	40.0	35.0	13.8	8.8	8.8	13.8	0.001	0.04	0.006	0.0043	2 799	220	3 019	2 045	344	2 389
12	HONOLULU, Hawaii	5 470	3.0	51.2	35.0	35.0	10.6	6.6	6.6	10.6	0.004	0.18	0.0125	0.013	847	35	882	506	58	564
13	TOKYO, Japan	9 560	2.5	42.5	40.0	35.0	7.0	2.0	2.0	7.0	0.03	0.31	0.036	0.024	193	5	198	63	21	84
14	LOWER HUTT, New Zealand	12 650	2.0	41.6	35.0	40.0	6.8	6.8	6.8	6.8	0.05	0.44	0.026	0.039	135	14	149	135	14	149
15	BYRD LAND, Antarctica	14 430	3.0	24.2	30.0	32.0	0/(1)	9.2	9.2	0/(1)	0.08	0.52	0.069	0.06	11	16	27	114	1	115
16	JOHANNESBURG, So. Afr.	16 150	4.5	17.6	35.0	48.0	3.7	2.7	2.7	3.7	0.1	0.6	0.17	0.45	20	3	23	5	5	10
17	BUENOS AIRES, Argentina	9 940	4.5	22.5	35.0	45.0	9.0	8.6	8.6	9.0	0.03	0.34	0.17	0.34	51	24	75	24	25	49

1. "NUMBER OF MEASUREMENTS" REPRESENT THE NUMBER OF FREQUENCY COMPARISON MEASUREMENTS POSSIBLE DURING THE TOTAL TIME OF COMPLETE DAYLIGHT AND COMPLETE DARKNESS BETWEEN BOULDER AND THE "LOCATION OF MEASUREMENTS."
2. PRECISION OF EACH FREQUENCY COMPARISON MEASUREMENT – 1 PART IN 10^9.
3. HIGHER PRECISION THAN 1 PART IN 10^9 WILL BE OBTAINABLE BY MAKING MEASUREMENTS FOR PERIODS LONGER THAN THE MINIMUM TIMES LISTED. OVER PATHS WHERE DAYLIGHT IS PRESENT FOR 2 HOURS OR MORE, PRECISIONS RANGING FROM 3 PARTS IN 10^{10} TO POSSIBLY 3 PARTS IN 10^{11} ARE EXPECTED

measurements with a precision of 1 part in 10^9 obtainable per 24 hr for 99 percent of all hours have been obtained by dividing the number of hours of all-day or all-night over the particular path by the appropriate minimum observing times and adding, as shown in columns 16 and 17 for the months of June and December. It can be seen from this tabulation that the service provided at various locations is appreciably different; however, it is interesting to observe that in all but one case an appreciable number of measurements can be made throughout each day, and that in the worst case, for transmissions to New Delhi, at least one comparison with a precision of 1 part in 10^9 is expected in each day. In many of the closer more favorable locations, it would appear that precision of 3 parts in 10^{10} or better can be expected where observing times of at least 2 hr or more over daylight paths are possible. The maximum precision obtainable that can be expected and any locations can readily be calculated using the procedures outlined earlier [Watt and Plush, 1959].

5. Appendix

5.1. Timing Accuracy Obtainable With an On-Off Keyed Carrier in the Presence of Thermal Noise

It is well known that the slope of the envelope at the one-half amplitude point of a keyed carrier passed through a filter with a 6 db bandwidth B_t is

$$\frac{dv}{dt} \simeq EB_t \tag{A1}$$

where E is the maximum or locked keyed carrier voltage, and dv/dt is the slope observed at the one-half carrier envelope amplitude point.

When thermal type noise is present, the envelope of the keyed carrier can be expected to have a standard deviation in amplitude of $\sigma(a)$ for $C/N \gg 1$ with an envelope detector or for all C/N ratios if a synchronous detector is employed. If the point at which the noise free keyed carrier crosses the $E/2$ point is chosen as a reference T_0, it can be seen that the actual crossing point in the presence of noise will vary about this point. The standard deviation of this crossing time is defined as $\sigma(T_e)$, and it is easily seen that

$$\sigma(a)/\sigma(T_e) = dv/dt = EB_t \tag{A2}$$

$$\sigma(T_e) = \sigma(a)/EB_t \tag{A3}$$

Since $E = \sqrt{2}\ C$ where C is the rms carrier and $\sigma(a) = \sqrt{2}\ N$ where N is the rms noise in the receiver IF band pass filter, we can write

$$\sigma(T_e) = \frac{1}{(C/N)B_t} \tag{A4}$$

When the receiver effective bandwidth is B_r(c/s) and we use the convention of $C/N_{1\ kc}$, i.e., the rms carrier to rms noise in a one kc/s effective band, we

obtain for a single sample

$$\sigma(T_e)=\frac{\sqrt{B_r/1{,}000}}{(C/N_{1kc})B_t} \qquad \text{(A5)}$$

The authors have benefited from many helpful discussions with numerous individuals. In particular, we would like to acknowledge information and suggestions obtained from J. R. Wait, J. A. Pierce, A. G. Jean, R. C. Kirby, W. Markowitz, K. A. Norton, J. M. Richardson, W. D. George, D. H. Andrews, and D. D. Crombie. The assistance of Mrs. Winifred Mau in the preparation of this manuscript is also gratefully acknowledged.

6. References

Allan, A. H., D. D. Crombie, and W. A. Penton, Frequency variations in New Zealand of 16 kc/s transmissions from GBR Rugby, Nature **177,** 178 (1956).

Budden, K. G., The reflection of very low frequency radio waves at the surface of a sharply bounded ionosphere with superimposed magnetic field, Phil. Mag. **42,** 833–851 (1951).

Budden, K. G., The propagation of a radio atmospheric—II, Phil. Mag. **43,** 1179–1201 (1952).

Bullard, E. C., Definition of the second of time, Nature **176,** 282 (Aug. 1955).

Casselman, C. J., D.P. Heritage, and M. L. Tibbals, VLF propagation measurements for the Radux-Omega navigation system, Proc. IRE **47,** No. 5, 829–839 (May 1959).

CCIR VIII Plenary Assembly, Vol. I, Report 66, p. 362 (Warsaw, 1956).

Chilton, Charles J., VLF phase perturbation associated with meteor shower ionization, J. Geophys. Research **66,** No. 2, 379–383 (Feb. 1961).

Crichlow, W. Q., D. F. Smith, R. N. Morton, and W. R. Corliss, World-wide radio noise levels expected in the frequency band from 10 kc to 100 Mc, U.S. Dept. of Commerce, NBS Circ. 557 (Aug. 1955). (U.S. Govt. Printing Office, Washington 25, D.C.) See ITU, Geneva (1957).

Crombie, D. D., Differences between east-west and west-east propagation of VLF signals over long distances, J. Atmospheric and Terrest. Phys. **12,** 110–117 (1958).

Doherty, R. H., G. Hefley, R. F. Linfield, Timing potentials of Loran-C, to be published in Proc. IRE (1961).

Jean, A. G., W. L. Taylor, and J. R. Wait, VLF phase characteristics deduced from atmospheric waveforms, J. Geophys. Research **65,** 907–912 (March 1960).

Lyons, H., Spectral lines as frequency standards, Annals, N.Y. Academy of Sciences, p. 831 (May 1952).

Mach, Ernst, Science of mechanics, 5th English ed., p. 272 (Open Court Publishing Co., LaSalle, Ill., 1942).

Markowitz, William, Astronomical and atomic times, U.S. Naval Observatory, Washington 25, D.C. (9 March 1959). See also: Time measurement, Encyclopedia Britannica **22,** 224–228 (1959).

Markowitz, W., R. G. Hall , L. Essen, and J. V. L. Parry, Frequency of cesium in terms of ephemeris time, Phys. Rev. Letters **1,** No. 3, 105–107 (Aug. 1, 1958).

Morgan, A. H., Proposal for a new method of time signal modulation of VLF carriers (to be published as an NBS Tech. Note, 1961).

National Bureau of Standards, Boulder Laboratories, Standard frequency and time signals, WWV and WWVH, NBS Misc. Publ. 236 (1960).

National Bureau of Standards, Boulder Laboratories, National standards of time and frequency in the United States, Proc. IRE **48,** No. 1, p. 105 (Jan. 1960).

Nyquist, H., and S. Brand, Measurement of phase distortion, Bell System Tech. J. **IX,** 522–543 (July 1930).

Pierce, J. A., Intercontinental frequency comparison by VLF radio transmission, Proc. IRE **45,** 794 (1957).

Pierce, J. A., Recent long distance frequency comparisons, IRE Trans. on Instrumentation **I–7,** pp. 207–210 (Dec. 1958).

Pierce, J. A., G. M. R. Winkler, and R. L. Corke, The 'GBR experiment': A trans-Atlantic frequency comparison between caesium-controlled oscillators, Nature **187,** 914–916 (July–September 1960).

Round, H. J., T. E. Eckersley, K. Tremellen, and F. C. Lunnon, Report of measurements made on signal strength at great distances during 1922 and 1923 by an expedition sent to Australia, J. Inst. Elec. Engrs. (London) **63,** No. 346, 933–1011 (Oct. 1925).

Steele, J. McA., The standard frequency monitor at the National Physical Laboratory, Proc. Inst. Elec. Engrs. (London), Part B, **102,** 155–165 (1955).

Stone, R. R., Jr., W. Markowitz, and R. G. Hall, Time and frequency synchronization of Navy VLF transmissions, IRE, Trans. on Instrumentation **I–9,** No. 2, 155–161 (Sept. 1960).

Wait, J. R., A study of VLF field strength data, both old and new, Geofisica pura e applicata **41,** 73–85 (1958/III).

Wait, J. R., Diurnal change of ionospheric heights deduced from phase velocity measurements at VLF, Proc. IRE **47,** 998 (1959).

Wait, J. R., A comparison between theoretical and experimental data on phase velocity of VLF radio waves, Proc. IRE **49,** 1089–1090 (June 1961).

Wait, J. R., A new approach to the mode theory of VLF propagation, J. Research NBS (Radio Prop. **65D,** No. 1, 37–46 (Jan.–Feb. 1961).

Wait, J. R., and K. P. Spies, A note on phase velocity of VLF radio waves, J. Geophys. Research **66,** 992–993 (March 1961).

Watt, A. D., and E. L. Maxwell, Measured statistical characteristics of VLF atmospheric radio noise, Proc. IRE **45,** 55–62 (Jan. 1957).

Watt, A. D., R. M. Coon, E. L. Maxwell, and R. W. Plush, Performance of some radio systems in the presence of thermal and atmospheric noise, Proc. IRE **46,** No. 12, 1914–1923 (Dec. 1958).

Watt, A. D., and R. W. Plush, Power requirements and choice of an optimum frequency for a world-wide standard-frequency broadcasting station, J. Research NBS (Radio Prop.) **63D,** No. 1, 35–44 (July–Aug. 1959).

(Paper 65D6—164)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Design of Panoramic Ionospheric Recorders

L. H. Heisler and L. D. Wilson

Contribution from Radio Research Board Laboratory, Electrical Engineering School, University of Sydney, Australia

(Received January 3, 1961; revised May 26, 1961)

The design of ionosondes and the many electronic techniques used in the panoramic type are considered. A complete ionosonde which is readily transportable, economic to construct, and has the additional facility of electronic scan over any portion of the frequency range 1.0 to 20 megacycles per second is described briefly.

1. Introduction

It is becoming increasingly obvious that one of the most valuable ground based instruments for upper atmosphere research is the ionosonde recorder. Electronic computers and simplified reduction techniques such as those of Kelso [1954], or more recently Schmerling and Ventrice [1959], enable electron density distributions to be readily obtained from the corresponding h'f records, and recent efforts to tabulate these results on a worldwide basis are providing invaluable information on global ionospheric morphology. The need for more frequent ionospheric samplings is suggested by the presence of traveling ionospheric disturbances described by Munro [1950] and Heisler [1958]. This work indicates that the ionosphere is a dynamic medium, in which almost continuous changes in ion density are occurring. Furthermore, recent discoveries of a traveling disturbance—E_s relationship by Heisler and Whitehead [1960] and the sporadic nature of E_s itself, suggest that spaced ionosonde recorders are essential to study growth of E_s patches and mechanics of ionospheric disturbances.

Such recorders are, of necessity, very complex, and involve many diverse electronic techniques. The panoramic type recorder as originally described by Sulzer [1949] is now invariably used in modified forms such as that due to Heisler [1955], or as in existing American type C4 recorders [Brown, 1959]. However, the equipment is usually bulky and expensive. Also there seems to be lack of detailed design information available on this type of recorder.

This paper considers ionosonde characteristics necessary to satisfy recent research requirements. Design principles of the panoramic type ionosonde are discussed in detail and finally a new recorder is described which is readily transportable and economic to construct, possessing the additional facility of electronic scan over any portion of the frequency range 1 to 20 Mc/s.

2. Basic Description

The basic ionosonde recorder consists of a receiver tuned synchronously to a transmitter which scans over the frequency range f_0 to f_1 Mc/s in the scan period T sec, at the scan rate of S scans per hour. The transmitter is pulsed, the pulse length being t μsec at a repetition frequency of F c/s, where F usually is equal to the local a-c supply frequency. The directly received pulse and echoes are displayed on one axis of a cathode ray oscillograph display calibrated in height, and a calibrated frequency scale is provided either by movement of the height scale across the axis of movement of the recording film or by movement of the film itself.

In older types of recorders synchronization between transmitter and receiver was maintained by the use of cams and band switching. In the panoramic type recorder this is avoided by covering the frequency range in a single sweep. This is achieved by heterodyning an oscillator continuously variable from f_0+f_x Mc/s to f_1+f_x Mc/s with an f_x Mc/s pulsed oscillator and amplifying the resultant f_0 to f_1 Mc/s signal to provide the transmitter output.

The same variable oscillator is used as the local oscillator for a superheterodyne receiver of intermediate frequency f_x Mc/s. This receiver has an untuned input stage coupled directly to aerial, and is thus tuned at any instant to the transmitter frequency insuring automatic tracking.

3. Design Requirements

The various factors which must be considered in arriving at particular design features are considered below:

3.1. Frequency Range

The upper frequency limit is normally imposed by the maximum value of F_2 critical frequency. At sunspot maximum this may exceed 15 Mc/s but is seldom greater than 20 Mc/s. This limit may be exceeded by the critical frequency of sporadic E but this is exceptional and current research practice is to consider percentage of time f_0E_s exceeds 5 Mc/s.

Several factors influence the selection of a lower frequency limit. At very low frequencies elaborate transmitter and receiver antenna systems are necessary due to radiation inefficiency, and at frequencies below 1.5 Mc/s broadcast station interference in many areas causes considerable lack of resolution on the reproduced ionogram.

Many ionospheric research projects require detailed examination of a small portion of the frequency spectrum. It is desirable therefore, that the frequency range be variable within the limits discussed above. This also readily enables expansion

or contraction of frequency range to accommodate sunspot maximum or sunspot minimum conditions.

3.2. Scan Period and Scan Rate

There is now reason to believe that almost continuous changes in ion density are occurring in the ionosphere, particularly in winter months when daylight records taken at intervals of 1 min often show marked differences from each other [Heisler, 1958]. In order, therefore, to prevent distortion and consequent misinterpretation of the record, particularly in deducing electron density profiles, scan period should be no greater than 15 sec.

Consequent deterioration of integrated signal noise ratio as indicated by Dieminger [1959] does occur but should preferably be improved by means other than increasing scan period, such as careful receiver design or by photographic integration as discussed in section 4.4.

The scan rate should be variable to allow several scans per minute in special studies e.g., sporadic-E and a low rate for economical normal use.

3.3. Scan Law

The increasing use of ionograms in electron density profile analysis demands higher scaling accuracies than previously, particularly in regions where ion density changes slowly with height. Irrespective of accuracy of individual frequency calibration points therefore, interpolation between these points must be accurate. King [1960] has stressed the value of a logarithmic law scan in certain simplified $N(h)$ reduction techniques, and this type of law is also of value in ionospheric prediction analysis.

However, in practice the usual logarithmic law on many ionograms is only approximate and error can occur in interpolating between frequency marks. Interpolation is expedited by use of a linear scan law and small variations in this will usually introduce errors of less magnitude than those due to difficulties in interpolating an approximate logarithmic scale. One feature of log scan law is the open frequency scale at low frequencies where considerable accuracy is often required. However, this is usually at the expense of accuracy in reading f_0F_2.

3.4. Transmitter Pulse Length and Receiver Bandwidth

Since minimum detectable signal power is inversely proportional to pulse length [Herb and Sinsheimer, 1947], pulse lengths should be as long as possible consistent with required resolution. Price [1960] suggests that there is little advantage in pulse length t being shorter than 100 μS due to loss of resolution by ionospheric dispersion. This argument, however, does not apply to sporadic E ionization where a pulse length of 35 μS would be required to resolve the 5 km stratification observed by Briggs [1958].

Pulse length is also a factor in determining receiver bandwidth and the optimum value is $1.2/t$ Mc/s [Purcell, 1947].

4. Design Principles

The following design principles apply mainly to panoramic ionosondes. For purpose of detailed discussion the recorder may be considered to consist of the following main sections: variable frequency oscillator, fixed frequency oscillator, transmitter, receiver, and calibration system.

4.1. Variable Frequency Oscillator (VFO)

This oscillator provides a substantially constant output over the frequency range (f_0+f_x) Mc/s to (f_1+f_x) Mc/s in the scan period T sec. Since the wide frequency spectrum output supplies energy to three individual circuits, namely transmitter mixer, receiver mixer, and frequency calibrator it is essential that oscillator output impedance be low. High voltage output is not necessary as receiver and calibrator require only a few volts for satisfactory operation, and for reasons discussed later in section 4.3a high level mixing in the transmitter mixer stage is difficult to achieve.

The rate of change of frequency $(f+f_x)$ Mc/s of the oscillator with time t determines the scan law of the recorder, and this law is usually logarithmic or linear in form. If a rotating capacitor is used to vary the oscillator frequency it is possible to shape the rotor plate so as to give the desired rate of charge of capacitance with rotation angle. In practice, however, only an approximation to the required law is obtainable due to unpredictable edge effects between rotor and stator, and an exact law necessitates use of a rotating cam profile with a cam follower mounted on the capacitor shaft.

Any suitable rotating capacitor may be made to give a similar approximation to the required law by applying three point tracking techniques [Green, 1943] by addition of a series capacitor.

Electronic control of scan frequency may be obtained over a limited range by use of the recently developed voltage variable capacitor or Varicap [Hammerslag, 1959] or by application of a ring type resistance capacity feedback oscillator developed by Cormack [1951]. The frequency of this oscillator can be varied over a considerable range by application of a control voltage.

Electronic sweep control has the additional facility that any sweep range f_0 to f_1 can be obtained by suitable preset biasing arrangements while still maintaining scan law. A similar facility in mechanical scan arrangements can only be obtained by changing cam profiles. This form of control also removes the usual limitations on scan rates placed by mechanical sweep arrangements. Scans may be repeated as often as desirable, and minimum scan periods are limited mainly by possible receiver detuning before return of ionospheric echoes.

4.2. Fixed Frequency Oscillator (FFO)

This oscillates at a frequency of f_x Mc/s and since the highly sensitive receiver in the recorder has a similar intermediate frequency, it is essential that

the oscillator be pulse modulated by the transmitter keying pulse to prevent continuous receiver paralysis. It is extremely difficult even with most elaborate oscillator shielding to prevent leakage of signal into the receiver circuits.

The inherent accuracy of frequency calibration depends upon the accuracy of this oscillator. However the necessary requirements are not stringent and it is shown in section 5.5 that there is little advantage in having a fixed frequency oscillator accuracy greater than about one part in three thousand. This can be readily obtained with sufficient long term stability by means of a carefully constructed pulsed Hartley oscillator [Gamertsfelder and Holdam, 1949]. Pulse modulation of crystal oscillators is difficult and it is usually necessary to crystal control an oscillator at a submultiple frequency f_x/n Mc/s and multiply this by pulsed amplifiers to give the required f_x Mc/s output. Unfortunately there may be leakage of continuous f_x Mc/s harmonic radiation into receiver circuits.

4.3. Transmitter

This consists essentially of a mixer to obtain the sweep frequency from the variable frequency oscillator output and the pulsed output of the fixed frequency oscillator, followed by a wide band amplifier feeding the transmitter aerial. These require a balanced output and hence push-pull output is usually employed, with the additional advantage of providing higher power output particularly in class B operation. Adoption of this necessitates some form of phase inversion in the amplifier chain to provide out of phase driving voltages for the output stage. In order to obtain sufficient gain with a minimum of tubes, wide band coupling techniques must be used throughout mixer and amplifier circuits. The various approaches which can be made to these problems are considered below.

a. Resistance Capacitance Coupling

In this type of coupling the top frequency to be amplified f_1 is approximately equal to bandwidth B_n. If the amplifier consists of n stages and if B is the bandwidth of each stage

$$B_n=\frac{B}{\sqrt{n}} \tag{1}$$

The normal maximum plate resistor R in any stage is given by

$$R=\frac{1}{2\pi BC} \tag{2}$$

where C is the interstage shunt capacity.

R may be increased by a factor of 1.77 to 2.4 by the use of two and four terminal network coupling techniques, described by Walker and Wallman [1948]. The total interstage capacity C should be measured as strays are unknown and published values of tube capacitances are cold values and have been found to deviate considerably from hot tube values.

Some resistors have been found to change value considerably at high frequencies, and R for preference should be deposited carbon on a ceramic former.

It is possible using network coupling techniques to deliberately peak the high frequency gain of an individual stage [Walker and Wallman, 1948]. If this is done then bandwidth shrinkage as expressed by eq (1) is not so severe, and bandwidth requirements for an individual stage may be relaxed.

The limitations imposed by maximum power rating of tubes may be overcome by pulse switching stages for the duration of the transmitter pulse. Screen voltages may then be increased by a factor N with consequent increase in mutual conductance g_m to $\sqrt{N}g_m$ approximately. Moreover the effective grid base of a tube is lengthened by a factor N approximately so that higher permissible grid voltage swings become possible, an important feature in high level stages.

One immediate consequence of pulse switching is that a large negative pulse appears across plate resistor R due to change in plate potential level when the valve becomes operative, and small coupling time constants consistent with adequate coupling at frequency f_0 are necessary to prevent appearance of this on the grid of the following stage.

The valve chosen as transmitter mixer must have a high value of g_m since conversion transconductance is approximately $g_m/3$. This restricts choice of mixer valves to those with a relatively short grid base. Consequently high voltage inputs are not possible and high level mixing is difficult to achieve.

b. Wide Band Transformer Coupling

One difficult problem associated with the transmitter design is to provide sufficient power output into a low impedance. In practice the push-pull output stage is directly coupled to the aerial circuits to develop power in an effective load of 600 ohms under class A operating conditions. The type of tube chosen as final amplifier must therefore have high mutual conductance, low plate resistance and large negative grid voltage for plate current cutoff, conditions which are most readily met by pulse modulator tetrodes specially developed for radar purposes.

Performance can be improved by matching aerial impedence to the plate resistance of the final amplifiers by means of a wide band transformer. This enables class B operation to be used with its much higher plate efficiency and provided sufficient power is available from the previous stage, grids may be driven positive, consequently valves need not have a high value of grid voltage cutoff. This relaxation in required characteristics permits use of normal high perveance valves and effects considerable economy.

A complete discussion of very wide band transformer design is given by Maurice and Minns [1947]. The upper frequency limit is mainly controlled by leakage inductance which is minimized by keeping number of turns low and careful attention to physical design. At low frequencies number of turns must

be sufficient to provide required inductance, and is reduced by using a core of modern powdered iron dust materials possessing high initial permeability. However, care must be taken to insure that flux saturation does not occur at the power outputs considered. If the transformer has balanced windings, then a high degree of balance can only be obtained by inserting an electrostatic shield between windings.

c. Phase Inversion

If the transmitter has push-pull output, some form of phase inversion is necessary in the amplifier chain to provide driving voltages 180° out of phase. If the phase splitter is used in the early stages of the amplifier then all subsequent stages must be push pull in form, with consequent duplication of valves, increase in power consumption and the risk of voltage unbalance occurring at the output stage grids. When the phase splitter is placed in late stages of the amplifier, it may be necessary to employ large pulse tetrodes.

A convenient means of obtaining the necessary inversion is by using an amplifier having unity gain as described by Heisler [1955], or the cathode coupled unbalance to balance system used in the C4 recorder [Brown, 1959].

Another method of realizing the necessary anti-phase voltages is by use of a broad band transformer with a balanced secondary winding. This can be used to advantage as a means of coupling the driver stage to power output stages operated under class B conditions.

d. Transmitter Pulser

This applies a large pulse of suitable polarity to either grid or screen circuits rendering the valve operative for a period of t μsec. During this period the pulser must present a low impedance to the appropriate circuits of the controlled valves. This permits loading by complex impedences in pulsed grid circuits and enables pulsed screen circuits to be operated without RF bypass capacitors, both factors impairing pulse shape, and inhibiting feedback between common pulsed stages.

In practice the bootstrap type pulser [Glasoe, 1948] pulsing screen grids has been found most satisfactory.

e. Transmitter Aerial

Effective power output of the transmitter depends to a large extent on the radiation efficiency of the associated antenna system. The broadband requirements demand design compromise particularly at the low frequency end, where physical dimensions of the antenna have to be at least a quarter of a wavelength for efficient radiation. Usually a 600-ohm delta type antenna is employed similar to that described by Cones et al. [1950]. Unfortunately, considerable power is dissipated in the terminating resistance and at low frequencies as much as 80 percent of the power input to the aerial may be lost in this manner [Cones, 1951]. Where research is concerned with this end of the spectrum, it may be necessary to provide high power input to the aerial system to compensate for this.

There appears to be considerable need for development of suitable aerial designs for this purpose, particularly with a view to maintaining aerial efficiency at low frequencies.

4.4. Receiver

This consists of an untuned aerial stage and mixer followed by sufficient intermediate frequency amplification, to provide several volts of signal at the detector with high signal noise ratio. Since all frequencies arriving at the aerial appear on the mixer grid, it is necessary to attenuate frequencies outside the signal band f_0 to f_1 Mc/s to prevent serious cross modulation. In practice a high pass filter is inserted in aerial circuits to remove broadcast station interference below f_0, while a resonant trap tuned to f_x Mc/s prevents breakthrough at the intermediate frequency of the receiver.

A similar restriction of pass band may be obtained by using a well designed wide band transformer on aerial input circuits. Aerial input is usually balanced in form and use of a balance to unbalance transformer greatly simplifies front end design.

Bandwidth requirements are mentioned in section 3.4. Intermediate frequency f_x is usually greater than 20 Mc/s, hence gain and selectivity can be improved by using double frequency conversion. As will be described later the frequency calibration system uses a 1 Mc/s harmonic generator; the value of second intermediate frequency should therefore be other than integral to prevent interference from this source.

The amplitude of the received pulses covers a large dynamic range and if receiver gain is adjusted to record weak echoes satisfactorily, strong echoes invariably saturate the receiver with consequent loss of complex splitting detail.

This effect may be minimized by some form of instantaneous automatic gain control [Lawson and Uhlenbeck, 1948a], or alternatively by some form of logarithmic receiver such as that described by Chambers and Page [1954], which has a logarithmic gain response over a very wide range of amplitude of input signals. The latter in particular is most suitable for measurement of absolute amplitudes in ionospheric absorption research applications.

Automatic gain control is particularly effective in discriminating against unwanted CW signals. Similar discrimination by differentiation of the received echo before application to display should be used with caution or loss of complex splitting detail in the received echo may occur. The usual RC filter is not satisfactory [Lawson and Uhlenbeck, 1950b] and a high pass filter with low frequency cutoff $1/(2t)$ should be used. The technique is only of real value if intermediate frequency amplifiers are not overloaded [Haworth and Jordon, 1947].

Receiver output should provide echoes of negative polarity for application to the grid of the cathode ray display tube. Photographic integration on the resultant ionogram has been shown to provide considerable noise discrimination.

4.5. Calibration System

The final ionosonde presentation must be provided with calibration of both frequency and height. Frequency calibration is usually provided by heterodyning the variable frequency oscillator output f_0+f_x to f_1+f_x Mc/s with harmonics from a 1 Mc/s harmonic generator and using the resultant zero beats to generate short sharp pulses corresponding to each megacycle in the range f_0 to f_1. These pulses are then applied to intensifier circuits in such a way to provide suitable marks on the visual display.

The generator is usually crystal controlled and accuracy of individual marks depends mainly on accuracy of f_x. However, since the received echo falling on a given frequency mark includes a number of frequencies in the optimum receiver bandwidth, it is doubtful whether frequencies read to any greater accuracy than one half of this have any real significance. This demands a FFO stability of 1 part in 3,000.

Height calibration appears on the final display as a series of sharp marks corresponding usually to 50 km height separation. These are generated by a 3,000 cycle oscillator gated so that it only operates during each height display.

An upper limit to overall accuracy of height calibration of at least 1 km is imposed by errors in time measurement due to rise time of received echoes and variations of receiver delay with signal amplitude [Chance, 1949]. The maximum height range usually employed is 1,000 km, and therefore accuracy is limited to one part in 1,000. This order of stability is easily attained in a well designed RC oscillator.

5. Recorder Description

An ionosonde is described which covers the frequency range 1 to 20 Mc/s. This has been constructed in the laboratory of the Radio Research Board Sydney, based on many of the design techniques discussed above. The complete recorder weighs approximately 150 lbs and is contained in a 22 in. square cubicle which stands 34 in. high (without camera and camera boxes). The ionosonde features complete electronic operation, with a nominal scan period of 15 sec repeated at a scan rate of 1 per minute, with a linear or logarithmic scan law. However, scan period can be made as short as 2 sec and this can be repeated almost continuously. Nominal frequency range is 1 to 20 Mc/s but any part of this range may be selected by the setting of two controls. A manual control permits a manual scan to be made, or if necessary the ionosonde may be set to operate on any particular fixed frequency. Transmitter pulse length is variable between 40 and 100 μsec and the pulse peak power input to a 600 ohm balanced antenna averages 3 kw at a pulse repetition frequency of 50 c/s. Most of the power supplies are electronically voltage regulated. The complete recorder can be constructed for less than £800 (Australian).

Because of space considerations it is obviously impossible to publish complete circuit diagrams or to give detailed technical information. The following description therefore lists briefly the most important technical features of the equipment.

5.1. Variable Frequency Oscillator (VFO)

This is developed from a circuit due to Cormack [1951]. Basically the circuit is a resistance capacitance ring type oscillator, in which the necessary phase shift is provided by a triode amplifier and four cathode followers. Phase rotation and hence oscillator frequency is controlled by varying the mutual conductance of the cathode followers with an applied voltage. The voltage developed by the oscillator is from 1.5 to 1.0 volts rms over the frequency range 29 to 48 Mc/s and the almost linear frequency range shift for applied control obtained is amplified to supply 4.0 to 6.0 volts rms to the transmitter and receiver mixers. The transmitter mixer is supplied directly from this amplifier while the receiver mixer and frequency calibrator mixer are coupled to the amplifier by cathode followers.

5.2. Fixed Frequency Oscillator (FFO)

This is a pulsed Hartley oscillator as described by Gamertsfelder and Holdam [1949a] in their figure 4.45. The oscillator tube operates over a linear region of the tube characteristics and is therefore free from frequency deviation effects due to stray power amplifier pulses in the grid circuits. C class pulsed oscillators are particularly prone to this. A buffer amplifier is used after the oscillator to drive the transmitter mixer cathode.

5.3. Transmitter

A type 12BY7 tube was chosen for the transmitter mixer from several types tried. The VFO voltage is injected into the control grid and the FFO voltage into the cathode circuit. The plate circuit is coupled to a 6CK6 amplifier by a four terminal network with a cutoff frequency of 22 Mc/s, which inhibits higher order component frequencies from reaching the 6CK6 grid. This amplifier is similarly coupled to a type 6CM5 amplifier which is transformer coupled to the power amplifiers. The transformer is a wide bandwidth type wound on a toroid A4 Ferroxcube core, providing out of phase voltages to drive the final push-pull amplifiers. These are type 6CM5 tubes coupled to antenna circuits by another wide bandwidth transformer. All amplifiers are screen grid pulsed by a common bootstrap type pulser unit.

5.4. Receiver

Type 6BX6 tubes are used as RF amplifiers and these are coupled to type 6EJ7 tubes as mixers by four terminal coupling networks. The variable oscillator voltage of average value 1.8 v rms is

injected into the cathode circuits of the mixer tubes. The mixer stage is followed by a single stage of IF amplification at 28 Mc/s which is then converted to 1.9 Mc/s and amplified by two additional stages. Dynamic range is improved by use of variable MU tubes throughout the amplifier and application of automatic gain control to the last stage. This is followed by a conventional compensated diode detector and one stage of video and audio amplification. The receiver has an overall gain of 114 db and a bandwidth of 30 kc/s.

5.5. Calibration System

This provides suitable range marks and frequency marks on the final display. The height calibrator uses a gated resistance capacity 15 kc/s oscillator and this signal is divided to provide a 3 kc/s signal. The 15 kc/s and 3 kc/s signals are clipped, suitably differentiated, mixed and applied to the display to produce fine 10 km and broad 50 km height marks. Either or both calibrations may be selected at will.

The frequency calibrator consists of a 6BL8 tube as a triode 1 Mc/s crystal oscillator and pentode harmonic generator. This output is mixed with the output of the variable frequency oscillator in the pentode section of a 6AJ8 tube and the resulting heterodyne beats which correspond to 1 Mc/s frequency intervals are amplified, squared and applied to a multivibrator which when operative, removes the time base synchronizing signal and hence the cathode ray tube display. Special provision is made to insure that frequency intervals always occupy the same deflection length on the display irrespective of duration of scan period, and a later improvement places frequency markers on top and bottom of the record only.

5.6. Control Unit

The control unit provides the basic variable d-c voltage to vary the frequency of the VFO. The same voltage is used to provide display tube deflection when a linear scan law is desired. An adjustable RC charging circuit with suitable controls to vary amplitude and initial starting voltage establishes start and finish of scan voltage and hence the particular frequency scan. A synchronous motor is used to provide regular 1 min or longer intervals between scans.

5.7. Display

Two displays are provided, one is on the face of a 5 in. long persistence screen cathode ray tube and is used for monitoring, while the second display is on a similar tube having a normal persistence screen which is photographed with either a 35 mm or 16 mm camera.

Typical positive prints of ionograms produced by the recorder are shown in figures 1, 2, and 3. Figures 1 and 2 are both examples of 12 sec sweeps. In figure 1 the frequency range covered is from 1 to 14 Mc/s, while in figure 2 the frequency range is 2 to 20 Mc/s. Figure 3 shows a very fast sweep of 2 sec over the frequency range 0.1 to 12 Mc/s. It will be noticed that the sweep is so rapid that individual strokes of the time base are obvious, yet in spite of this the amount of information available does not differ materially from that in figures 1 and 2. The examples serve to illustrate the great versatility of the equipment both in scan period and frequency range.

FIGURE 1. *Typical ionogram covering the frequency range 1.0 to 14 Mc/s in a scan period of 12 sec.*

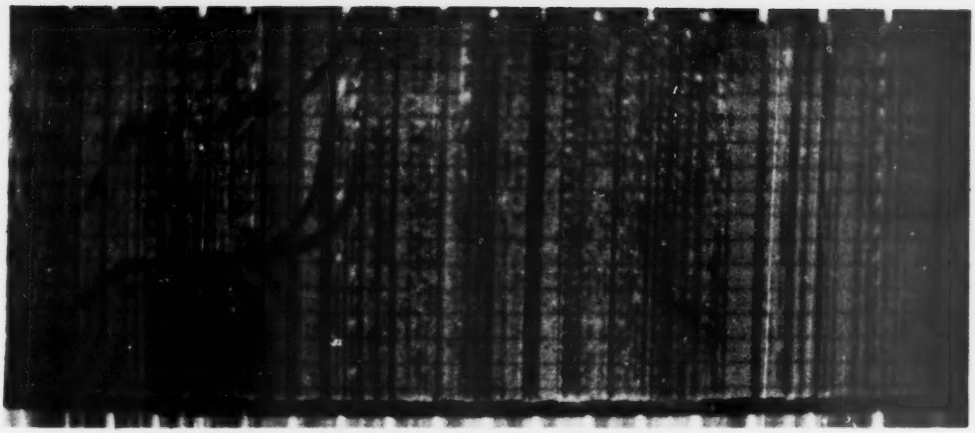

FIGURE 2. *Typical ionogram covering the frequency range 2.0 to 20 Mc/s in a scan period of 12 sec.*

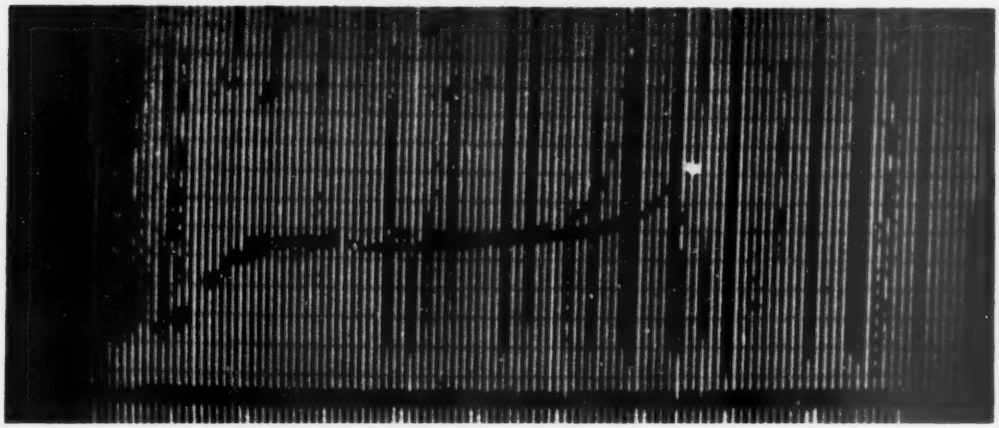

FIGURE 3. *Typical ionogram covering the frequency range 1.0 to 12 Mc/s in a scan period of 2 sec.*

6. Conclusions

Many facets of panoramic ionosonde design have been considered. There are possibly other design features, particularly for specific research projects which have not been included, but it is felt that the factors discussed constitute enough information to provide a basic ionosonde design. Some of the main design features emphasized are as follows:

(i) The use of electronic scan control provides many desirable ionosonde features economically. Rapid scans over the whole or any part of the frequency range are possible, and scan laws either logarithmic or linear can be selected at will.

(ii) Considerable economy can be effected in transmitter circuits by the use of wide band transformer coupling.

(iii) There is little advantage in using transmitter pulse lengths less than 100 μsec, except for E_s investigations where a pulse length of 40 μsec may be beneficial.

(iv) Present transmitter antenna systems are very inefficient at low frequencies. There is need for research into suitable aerials for this particular purpose.

(v) To avoid technical design difficulties, the frequency range of the ionosonde should not be greater than necessary. A range of 1.0 to 20 Mc/s is suggested.

(vi) The scan period should not exceed 15 sec.

(vii) The scan rate should be frequent at least one frame per minute, although this of necessity is governed by economy in film consumption.

(viii) Differentiation of receiver output is advantageous, only if intermediate fre-

quency stages are not overloaded and care is taken in requisite filter design.

(ix) Considerable noise discrimination is obtained if the signal applied to display cathode ray tube grid is of negative polarity.

(x) There is some doubt as to the advantage of a logarithmic scan law, and interpolation errors in such a case are likely to be greater than those occurring in a similar linear law. Electronic scan control is particularly adaptable to generation of a linear scan law.

(xi) For most purposes a receiver bandwidth of 30 kc/s is more than adequate even for scan periods of the order of a few seconds.

A complete technical description will be given in a later paper.

This work has been carried out in the Electrical Engineering School of the University of Sydney as part of the research program sponsored by the Radio Research Board of the Commonwealth Scientific and Industrial Research Organisation.

We express our appreciation to the University of Sydney and in particular to Professor W. N. Christiansen for provision of facilities in the Electrical Engineering School. The authors also thank Dr. G. H. Munro for helpful discussions during the preparation of the paper.

7. References

Briggs, B. H., Some characteristics of sporadic-*E* ionisation as determined by a rapid frequency sweep experiment, Agardograph 34, sporadic-*E* ionisation, 151–161 (NATO advisory group for aeronautical research and development, 1958).

Brown, J. N., Automatic sweep-frequency ionosphere recorder Model C–4, Proc. IRE **47,** No. 2, 296–300 (Feb. 1959).

Chambers, T. H. and I. H. Page, The high-accuracy logarithmic receiver, Proc. IRE **42,** No. 8, 1307–1314 (1954).

Chance, B., Electronic time measurements, 1st ed., Radiation Laboratory Series, Vol. 20, ch. 3, 39–41 (McGraw-Hill Book Co. Inc., New York, N.Y., 1949).

Cones, H. N., Effect of a metal mast and guy wires on the performance of a 600 ohm multiple-wire delta antenna, J. Research NBS **46,** No. 2, 113–120 (Feb. 1951).

Cones, H. N., H. V. Cottony, and J. M. Watts, A 600 ohm multiple-wire delta antenna for ionospheric studies, J. Research NBS **44,** 475–488 (May 1950).

Cormack, A., Wide range variable-frequency oscillator, Wireless Engr. **28,** No. 336, 266–270 (Sept. 1951).

Dieminger, W., Transient fine structure of the *E* layer, J. Atmospheric and Terrest. Phys. **16,** Nos. 1/2, 179 (Oct. 1959).

Gamertsfelder, G. T., and J. V. Holdam, Waveforms, 1st ed., Radiation Laboratory Series, Vol. 19, ch. 4, 143–145 (McGraw-Hill Book Co., Inc., New York, N.Y., 1949).

Glasoe, G. N., Pulse generators, 1st ed., Radiation Laboratory Series, Vol. 5, ch. 4, 120 (McGraw-Hill Book Co., Inc., New York, N.Y., 1948).

Green, A. L., Superheterodyne tracking charts III, The padded signal circuit, Wireless Engr. **20,** No. 243, 581–594 (Dec. 1943).

Hammerslag, J., Circuit design using silicon capacitors, Electronics **32,** No. 38, 48–50 (Sept. 18, 1959).

Haworth, L. J. and W. H. Jordon, Radar System Engineering, 1st ed., Radiation Laboratory Series, Vol. 1, ch. 12, 459 (McGraw-Hill Book Co. Inc., New York, N.Y., 1947).

Heisler, L. H., A panoramic ionospheric recorder for the study of travelling disturbances in the ionosphere, Australian J. Appl. Sci. **6,** No. 1, 1–6 (Jan. 1951).

Heisler, L. H., Anomalies in ionosonde records due to travelling ionospheric disturbances, Australian J. Phys. **11,** No. 1, 79–90 (Jan. 1958).

Heisler, L. H. and J. D. Whitehead, *F*-region travelling disturbances and sporadic-*E* ionisation, J. Geophys. Research **65,** No. 9, 2767–2773 (Sept. 1960).

Herb, R. G., and R. L. Sinsheimer, Radar system engineering, 1st ed., Radiation Laboratory Series, Vol. 1, ch. 15 596 (McGraw-Hill Book Co. Inc., New York, N.Y., 1947).

Kelso, J. M., The determination of the electron density distribution of an ionospheric layer in the presence of an external magnetic field, J. Atmospheric and Terrest. Phys. **5,** No. 1, 11–27 (Mar. 1954).

Kelso, J. M., The calculation of ionospheric electron density distributions, J. Atmospheric and Terrest. Phys. **10,** No. 2, 103–109 (Feb. 1957).

King, G. A. M., Use of logarithmic frequency spacing in ionogram analysis, J. Research NBS **64D,** No. 5, 501–504 (Sept.–Oct. 1960).

Lawson, J. L., and G. E. Uhlenbeck, (a) Threshold signals, 1st ed., Radiation Laboratory Series, Vol. 24, ch. 11, 303–306 (McGraw-Hill Book Co., Inc., New York, N.Y., 1950).

Lawson, J. L. and G. E. Uhlenbeck, (b) Threshold signals, 1st ed., Radiation Laboratory Series, Vol. 24, ch. 12, 342–343 (McGraw-Hill Book Company Inc., New York, N.Y., 1950).

Maurice, D., and R. H. Minns, Very wide band radio frequency transformers, Wireless Engineer **24,** Nos. 285 and 286, 168–177, 209–216 (June and July 1947).

Munro, G. H., Travelling disturbances in the ionosphere, Proc. Roy. Soc. **A202,** 208–223, 1950.

Price, W. L., (1960) private communication.

Purcell, E. M., Radar system engineering, 1st ed., Radiation Laboratory Series, Vol. 1, ch. 2, 33–35 (McGraw-Hill Book Co., Inc. New York, N.Y., 1947).

Schmerling, E. R., and C. A. Ventrice, Coefficients for the rapid reduction of *P'-f* records to *N-h* profiles without computing aids, J. Atmospheric and Terrest. Phys. **14,** Nos. 3/4, 249–261 (June 1959).

Sulzer, P. G., Sweep frequency ionosphere equipment, J. Appl. Phys. **20,** No. 2. 187–196 (Feb. 1949).

Walker, R. M., and H. Wallman, Vacuum tube amplifiers, 1st ed., Radiation Laboratory Series, Vol. 18, ch. 2, 75–79 (McGraw-Hill Book Co., Inc., New York, N.Y., 1948).

(Paper 65D6–165)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

A Quick Method for Estimating the Stage of the Sunspot Cycle

W. B. Chadwick

Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

(Received June 22, 1961)

A method, based on the maximum median hourly value of f_0F_2 month-by-month as observed at Washington, D.C., is given for predicting a smoothed annual sunspot number immediately at the close of a given month, centered on the month in question. Regression equations and standard errors are given. It should be a useful supplement to the McNish-Lincoln method of prediction of sunspot numbers, particularly during the first two years of the rising part of a sunspot cycle. It is capable of use with the observations from any ionosphere sounding station operated in a consistent manner over a period of years, preferably during at least two solar cycles.

1. Introduction

One of the widely accepted methods used in the prediction of sunspot numbers for use in radio frequency predictions is the McNish-Lincoln method [1949]. This method involves the use of smoothed annual sunspot numbers corrected by departures one year earlier than the time in question. It predicts month by month the "best" prediction for 1 year ahead.

The use of a 12-month running average suffers of course from the fact that to obtain the sunspot number applicable to a certain month, 6 months in the future must be observed. When, however, the authors extended the technique to the prediction of 3-month mean sunspot numbers, lower overall reliability resulted. The method outlined here has the advantage that, immediately upon the completion of a month, a prediction of the 12-month running average sunspot number centered on that month can be made. Of considerable importance is the fact that the rms deviation of prediction would seem to be of about the same order of magnitude as those given in a second paper [McNish and Lincoln, 1954] in which the authors examined the reliability of their method of prediction and applied it to cycle 18.

In the proposed method, the highest monthly median value of f_0F_2 for *any hour*, as observed at Washington, D.C., is used, in a set of 12 regression equations, one for each month, to obtain an estimate of the 12-month running average sunspot number for the month that has just ended. It is believed that this choice of maximum f_0F_2, rather than f_0F_2 for a fixed time of day, is of aid in minimizing seasonal effects.

The idea of using f_0F_2 as a measure of solar activity is not new, but the methods so far proposed have been more elaborate, involving for instance such things as 5-hourly averages centered on noon, and averaging of results from several stations [Phillips, 1947; Minnis, 1955]. Also much work has been done and is continuing on other possible indices [Kundu and Denisse, 1958; Minnis and Bazzard, 1959]. The fact that a reasonable answer can be obtained to the problem of predicting solar activity by the simple method outlined here, gives reason for hope that the much more elaborate methods presently under study will eventually result in a highly satisfactory index obtained from some combination of physical quantities. At the same time, it is distinctly worthwhile to have some way of connecting future ionospheric indices with the long time-series of sunspot numbers now available, on which so many relationships have previously been based.

The method of this paper is intended to show what can be done almost immediately using data from but one ionosphere station (Washington) which has been operated in a consistent fashion over a long period of time. The initial impetus for this study was the sudden drop in solar activity in January 1961 as represented by the provisional Zürich monthly sunspot number of 53.5, together with the desire to investigate what this sudden drop might mean in terms of future predictions of critical frequencies.

2. Correlation Coefficients

Various results are given in table 1. It will be noted that, in general, the coefficients in column A exceed those in column B, and those in column E exceed those in column D. (However, these observed tendencies by themselves are not statistically significant.) Column C is filled out only for January and May, the months for which the highest and lowest coefficients were obtained in column A. The omission of numbers for the maximum period 1957 to 1959 (col. C) and the reduction of high numbers in this period (col. E) gave correlations which, when taken with those in columns A and D, form an interesting series of coefficients for these months.

TABLE 1. *Correlation coefficients, maximum hourly median f_0F_2 (Washington) versus Zürich sunspot number for various periods; A, C, D, versus 12-month running average sunspot number centered on the month in question; B, versus monthly sunspot number; E, same as D, but with sunspot numbers >150 used as 150*

	A	B	C	D	E
	1945–1960	1945–1960	1941–1956, 1960	1941–1960	1941–1960
Jan	0.989	0.97	0.985	0.989	0.975
Feb	.98	.94		.98	.98
Mar	.97	.97		.96	.97
Apr	.93	.92		.94	.97
May	.88	.91	.91	.90	.93
June	.92	.89		.93	.96
July	.92	.93		.94	.96
Aug	.92	.89		.93	.95
Sep	.95	.95		.96	.98
Oct	.96	.93		.97	.97
Nov	.98	.94		.98	.98
Dec	.97	.94		.98	.98

The use of a value of 150 for all running average sunspot numbers over 150, produced an equal or better correlation, except for January. This is in agreement with previously noted results [Bennington, 1958; Ostrow, 1959] and with CRPL predictions practice during the recent maximum.[1] In January, at the recent maximum period, increased solar activity did seem to have a marked effect in raising the criticals at Washington to a very high value in the middle of the day, higher than CRPL predictions based on sunspot number 150. Also it will be noted that correlations are somewhat higher for the winter months, a not unexpected result, as it has long been noted at Washington that the range of variation in f_0F_2 over the solar cycle, is much greater for winter months than for summer months. While it might be argued that the apparent seasonal effect was due to the way the sunspot numbers happened to fall, e.g., if there had been four numbers for January around 200 and none for June greater than 150, it should be remarked that during the recent maximum (1956 to 1959) each month had one 12-month running average sunspot number >187, and each month had three >150. Column B is included to show that correlations can be very high between monthly sunspot numbers and critical frequency. Figure 1 illustrates the relationship between f and R leading to the correlation coefficient 0.97 in column E, table 1, for March.

[1] CRPL Series D, Basic Radio Propagation Predictions, U.S. Government Printing Office, Washington 25, D.C., Nos. 146-172. Predictions for Jan. 1957–Apr. 1959, issued Oct. 1956–Jan. 1959.

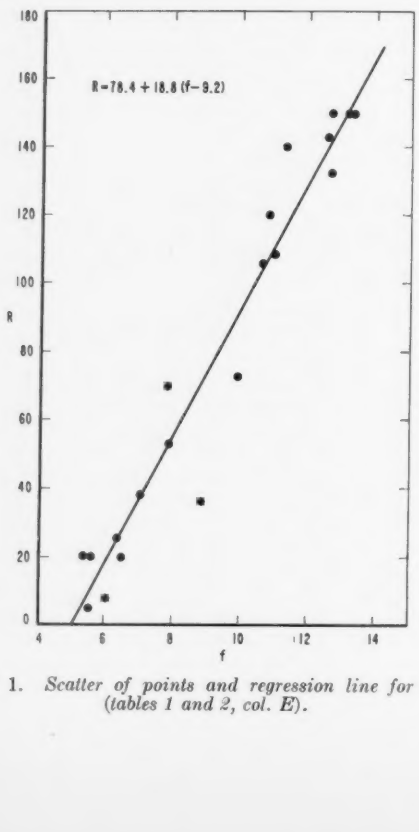

FIGURE 1. *Scatter of points and regression line for March (tables 1 and 2, col. E).*

3. Regression Equations

Table 2 gives the linear regression equations for each month of the year, based on the coefficients in column E of table 1. In addition, equations are given for October through February using the coefficients of column D, which are greater than or equal to the corresponding coefficients in column E.

TABLE 2. *Regression equations; R=12-month running average sunspot number, f=maximum hourly median f_0F_2(Mc/s); columns D, E correspond to D, E in table 1*

	D	E
Jan	$R=83.2+20.5(f-9.6)$	$R=78.4+17.8(f-9.6)$
Feb	$R=83.4+21.7(f-9.6)$	$R=78.4+19.0(f-9.6)$
Mar		$R=78.4+18.8(f-9.2)$
Apr		$R=78.8+24.1(f-8.2)$
May		$R=79.2+39.4(f-7.3)$
June		$R=79.5+56.0(f-6.8)$
July		$R=79.8+53.4(f-6.4)$
Aug		$R=80.0+44.1(f-6.8)$
Sep		$R=79.9+26.3(f-8.1)$
Oct	$R=84.2+22.9(f-10.0)$	$R=79.8+20.3(f-10.0)$
Nov	$R=84.2+21.4(f-10.3)$	$R=79.6+18.8(f-10.3)$
Dec	$R=84.0+22.9(f-10.0)$	$R=79.4+20.2(f-10.0)$

4. Standard Error of an Estimated Sunspot Number

Table 3 gives the low and high values of f (maximum hourly median f_0F_2 in Mc/s) month by month over the 20-year period 1941 to 1960, as well as the means, which, of course, are repeated from table 2. Table 4 gives the standard error of an estimated sunspot number, R, for each value of f given in table 3. These are computed from the formula

$$s_{R(\text{est.})}=s_{R\cdot f}\sqrt{1+\frac{1}{n}+\frac{(f-\bar{f})^2}{\sum(f_i-\bar{f})^2}}$$ [Snedecor, 1946]

where n is the number of observed pairs (f_i, R_i), $\bar{f}=\sum f_i/n$, $s_{R\cdot f}$ is the sample standard deviation about the regression line, and f is the predicting value of f_0F_2 as before. They may be compared with the

TABLE 3. *Range of values of f (Mc/s), Washington, 1941–1960*

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Low	5.9	5.7	5.4	5.3	5.2	5.3	4.8	5.0	5.2	6.2	6.5	6.2
Mean	9.6	9.6	9.2	8.2	7.3	6.8	6.4	6.8	8.1	10.0	10.3	10.0
High	15.5	14.0	13.2	11.2	10.0	8.0	7.7	8.6	11.0	14.2	14.8	14.0

TABLE 4. *Standard error of an estimated R for various values of f (see table 3)*

	Jan. Col. D, table 2	Jan. Col. E, table 2	Feb. Col. D, table 2	Feb. Col. E, table 2	Col. E, table 2							Oct. Col. D, table 2	Oct. Col. E, table 2	Nov. Col. D, table 2	Nov. Col. E, table 2	Dec. Col. D, table 2	Dec. Col. E, table 2
					Mar.	Apr.	May	June	July	Aug.	Sep.						
For low f	10.9	13.3	13.7	11.1	14.3	16.2	23.5	18.1	17.8	19.3	13.0	17.8	16.1	12.8	12.4	14.5	12.2
For mean f	10.6	12.8	13.1	10.6	13.8	15.5	22.1	17.1	16.8	18.3	12.4	17.0	15.4	12.3	11.9	13.8	11.6
For high f	11.7	14.1	13.9	11.3	14.4	16.2	23.8	17.7	17.5	19.3	13.0	18.0	16.3	13.0	12.6	14.6	12.3

rms deviation of prediction given in table 1 of McNish and Lincoln [1954], in which the deviations range from 7.0 to 19.2. Although these are of the same order of magnitude, it must be remembered that the McNish-Lincoln formulas were obtained using data from 10 solar cycles, whereas the standard errors of this paper were calculated using less than 2 cycles. Essentially the sunspot numbers used in the McNish-Lincoln method form a series of numbers spaced one solar cycle or about 11 years apart, whereas in this method the sunspot numbers are spaced 1 year apart.

5. Results and Conclusions

Since this paper was begun, values of maximum hourly median f_0F_2 at Washington have been received for January through May 1961. These are listed in table 5, together with the values of 12-month running average Zürich sunspot number estimated by the method of this paper, and those predicted by the McNish-Lincoln method. It must be remembered, of course, that the latter predictions were made, for the corresponding months, 6 months earlier than the predictions using the critical frequencies at Washington.

At the time of final version of this paper (June 5, 1961) the provisional 12-month running average Zürich sunspot number, centered on December 1, 1960, had dropped to 84.3. At the average rate of decline for the past 4 months, the number centered on the middle of January 1961 would be 76.7, compared with the predicted values 75.0 and 97.4 of table 5.

TABLE 5. *Values of f at Washington, estimated values of R, predicted values of R by McNish-Lincoln method, Jan.–May 1961*

1961	f (Mc/s)	R (Est.)	R (Pred., McNish-Lincoln)
Jan	9.2	[a] 75.0	97.4
Feb	8.6	59.4	93.8
Mar	8.6	67.1	90.2
Apr	7.4	59.5	88.5
May	6.8	59.5	86.2

[a] From table 2, col. D; all others from table 2, col. E.

A quick method of this type should prove especially valuable during the first 2 years of the rising part of a solar cycle, to help give an early indication of the level and timing of the coming maximum.

It would of course be desirable to augment the sample size for estimating the regression equations of table 2 as more years of data become available, in particular the years 1961 and 1962, thus completing a 22-year period. Also it should be noted that a similar thing could be done for any ionosphere sounding station having a long run of well-taken data. Thus delays in waiting for information from other parts of the world could be avoided.

An indication of the degree of accuracy at certain other locations may be obtained from table 6. Estimated values of R for January 1961 from the regression equations given in this table, are, for the respective stations, 70.6, 75.7, 81.4, and 84.5.

It is suggested that adequate results might be obtained from two locations, one in each hemisphere, at about the geomagnetic latitude of Washington, each station to cover its own winter months. For the

TABLE 6. *Correlation coefficients, regression equations, standard errors for four stations, 1945 to 1960, Jan., July*

Station	Geomag. lat.	Corr. coeff.		Regression equations		[b] Standard errors	
		Jan.	July	Jan.	July	Jan.	July
Fairbanks	64° N	0.98	0.91	$R=96.0+21.2(f-8.3)$	$R=98.5+64.4(f-5.7)$	13.9	27.7
Washington	50° N	.99	.92	$R=96.0+20.3(f-10.2)$	$R=98.5+60.7(f-6.7)$	9.5	24.3
Christchurch	48° S	.94	[a] .96	$R=96.0+48.7(f-7.7)$	[a] $R=102.4+22.8(f-8.9)$	22.2	[a] 18.9
Huancayo	1° S	.97	.96	$R=96.0+28.7(f-11.2)$	$R=98.5+35.7(f-8.9)$	15.8	17.3

[a] 15 years; data missing for July 1955.
[b] Standard error of an estimated R for mean f.

equinoxes, possibly the station with the longer time series would be used.

Finally, the presentation of two sets of equations in table 2 (columns D and E) is intended to be illustrative of two ways of obtaining the result. One should not necessarily switch from one type of equation to another, because in small samples such as these the reversal in size of correlation coefficients or other criteria could happen frequently and not be significant.

6. References

Bennington, T. W., Apparent saturation in $F2$ Layer, Wireless World **64,** No. 10, 472–473 (Oct. 1958).

Kundu, M. R., and J. F. Denisse, Solar radiation on decimeter waves as an index for ionospheric studies, JATP **13,** Nos. 1/2, 176–178 (July–Dec. 1958).

Minnis, C. M., A new index of solar activity based on ionospheric measurements, JATP **7,** No. 6, 310–321 (Dec. 1955).

Minnis, C. M., and G. H. Bazzard, Some indices of solar activity based on ionospheric and radio noise measurements, JATP **14,** Nos. 3/4, 213–228 (Mar.–Apr. 1959).

McNish, A. G., and J. V. Lincoln, Prediction of sunspot numbers, Trans. AGU **30,** No. 5, 673–685 (Oct. 1949).

McNish, A. G., and J. V. Lincoln, Predicition of the present sunspot cycle, Trans. AGU **35,** No. 5, 709–710 (Oct. 1954).

Ostrow, S. M., Variation of critical frequency with sunspot number, Paper presented at Spring Meeting of URSI, Washington, 1959.

Phillips, M. L., The ionosphere as a measure of solar activity, Terrestrial Magnetism and Atmospheric Elect. **52,** No. 3, 321–332 (Sept. 1947).

Snedecor, G. W., Statistical Methods, The Iowa State College Press, 137 (4th ed., 1946).

(Paper 65D6–166)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Measurements of Low-Angle Radiation From a Monopole

Alvin C. Wilson

Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

(Received April 27, 1961; revised June 27, 1961)

Experimental measurements using scale model techniques have been carried out to determine the effectiveness of a ground system of long-wire radials to obtain low angles of departure of transmission. Since transmission was to be in one direction only, the ground wires were laid out to form a ground-plane sector approximately 18° wide centered in the direction of transmission. The antenna was a base-driven vertical monopole. Measurements were made of the relative response in decibels for the monopole used as a receiving antenna at a frequency of 400 megacycles per second. The target transmitter antenna was always located at a distance of 200 wavelengths. At this separation the ground plane sector was in the near field of the target transmitting antenna and appropriate corrections must be made.

The received signal strength improvement due to the presence of the ground sector was approximately 14 decibels. The measured lobe positions of the first and second beam maximums and the first null are in good agreement with theory.

1. Introduction

In connection with a HF arctic propagation project using a base-driven vertical monopole at a frequency of 10 Mc/s over a ground of lava rock having a poor conductivity, a requirement arose to evaluate the effectiveness of a planned long-wire ground radial system in improving low-angle response.

A common practice has been to use a ground system of wire radials centered at the antenna base and extending outward for a distance of approximately one-half wavelength [Laport, 1952]. However, an improved ground system was desired to obtain low angles of departure of transmission.

One method of obtaining a low angle of departure would be to elevate the base of the antenna above ground. However, this would result in multiple lobes in the vertical plane. It was then believed to be advantageous to improve on the effectiveness of the ground radial system and maintain the antenna base at ground level.

Since the transmission path was in one direction only, a ground system was planned which would consist of wire radials 3,000 m long (approximately 100 wavelengths) covering an approximate segment of ground 18° wide in the desired direction of transmission.

Before the start of any actual construction, a study of this ground system of long-wire radials, using modeling techniques, was desired. Measurements were made to obtain the signal response over the center of the ground radial system and at other points over the ground-plane sector compared to the signal response over the ground path only. This would show the effect of the ground-plane sector width, the signal improvement due to the long wire radials, and the improvement in response with different numbers of wire radials (density of wires in the ground-plane sector).

This report presents the results of the measurements which were carried out at a frequency of 400 Mc/s using scaled model techniques.

2. Method of Measurement

For the purposes of measurement, the ground radial system to be measured was scaled for a frequency of 400 Mc/s (a scaling factor of 40). The measurements were carried out at the NBS antenna range on Table Mesa near Boulder, Colo. The ground area over which the measurements were carried out was first cleared of weeds, brush, and large rocks in order to make the ground as smooth as possible. The ground conductivity is considered to be poor and no attempt was made to scale the ground conductivity.

A quarter wavelength monopole, extended vertically from the center of a flat aluminum disk having a diameter of one wavelength. This disk was to represent a conventional ground system of a number of wires each one-half wavelength long.

The long-wire ground system initially consisted of 97 ground wires of No. 22 AWG Copperweld, each 75 m long. These wires covered an approximate 18° sector. The ends of the wires, separated by $\lambda/3$(0.191° angular separation) were secured in position and bonded to 20 cm-long steel spikes which were driven into the ground until the tops of the spikes were flush with the ground surface. The other ends of the radials were bonded to the edge of the flat aluminum disk.

The vertical monopole was used as a receiving antenna. The target transmitter antenna, a 5-

element Yagi oriented for vertical polarization, was moved in a 40° azimuthal arc at a distance of 75 m in front of the wire radial ends. The whole ground plane sector was not in the far field of the transmitting Yagi antenna. The measurements were made by recording the relative received response in decibels by the monopole. Figure 1 is a plan view of the ground radial system. The response in decibels versus azimuth position of the transmitting antenna was measured at a constant angle of arrival of 1.1°. The response in decibels versus angle of arrival in the range of 0° to 6.73° was measured at the sector center position. The angle of arrival is the included angle at the base of the monopole between the horizon and the target transmitter antenna.

The measurements of response in decibels versus azimuth position of the transmitting antenna at a constant angle of arrival of 1.1° were also carried out for ground systems of 49 wires (0.382° angular separation) and 25 wires (0.764° angular separation) by removing every other wire from the original 97 and then remaining 49 wires.

3. Results

Figure 2 presents the curves of recorded relative response in decibels versus azimuth position of the transmitting antenna for the ground system consisting of 97, 49, and 25 wires, measured over a 40° ground sector. The received signal is generally higher when measured at azimuth positions over the ground-plane sector center as is to be expected. The received signal is also highest in the case when the ground-plane sector was of maximum density of wire radials (97 wires). The fall-off of the received signal when measured over a path of ground only (no long wire radials) is gradual, and at 20° each side of the ground-sector center the long wire radials do not appear to have any effect on the received signal strength.

Figure 3 is a plot of the relative response in decibels versus the angle of arrival in the forward direction with a ground-plane sector of 97 wires. Measurements at angies above 6.73° were not possible because of the maximum height of mounting for the target transmitter antenna and separation distance between the receiving monopole and transmitting antenna.

4. Discussion of Results

On the assumption that the wire radials form a well conducting ground plane, and from the geometry of figure 1, computations can be made to compare theoretically computed values [Wait and Conda, 1958a] with the experimental results shown in figure 3.

The computed position of the first beam maximum is 3.3°. The experimental value of this first maximum is 2.4°. The position of the second maximum is computed to be 6.2° and the experimental value

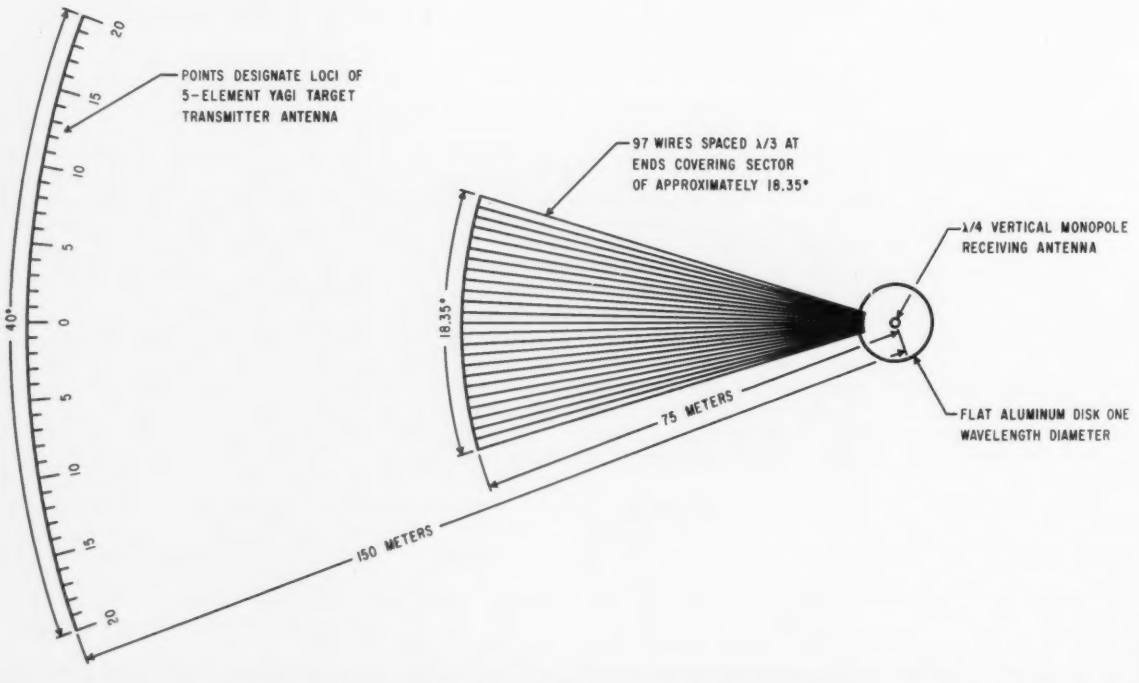

FIGURE 1. *Plan view drawing of extended wire radial ground system modeled for measurement at 400 Mc/s.*

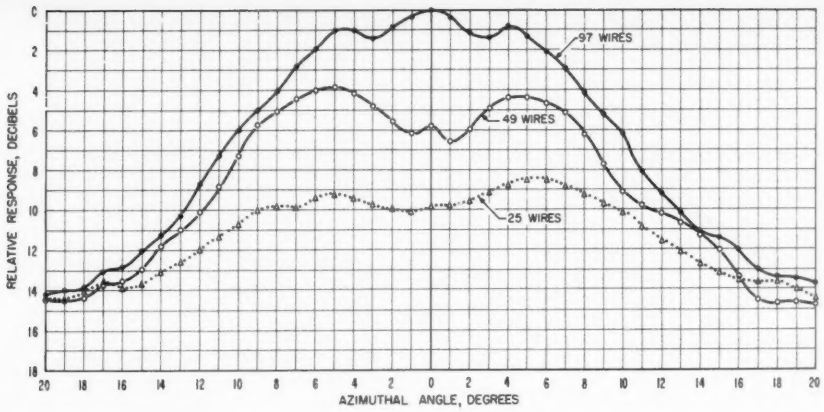

FIGURE 2. *Curves of relative response, decibels, versus azimuthal angle, degrees, for extended wire radial ground systems of 97, 49, and 25 wires.*

The angle of signal arrival is 1.10 in all cases.

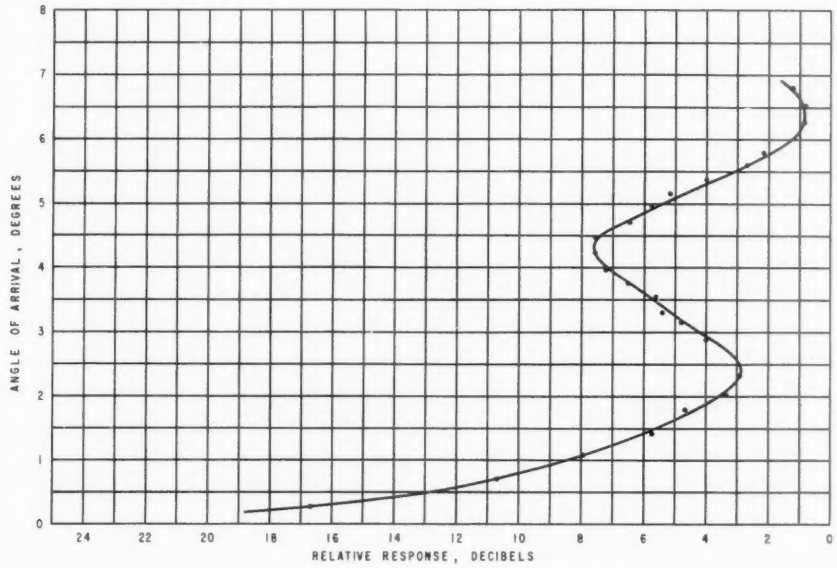

FIGURE 3. *Relative response of the monopole in the vertical plane in the forward direction over an extended wire radial system of 97 wires.*

is 6.3°. The position of the experimental first null is at an angle of 4.3°. This compares to a computed first null position of 5.0°.

With the exception of the first maximum, these experimental results agree quite well with a simple model of conducting half plane to represent the ground plane. The finite conductivity of the ground outside the ground plane will tend to lower the first lobe maximum [Wait and Conda, 1958a].

While the results of the model measurements are optimistic as regards low angle radiation, special considerations must be made when interpreting the results of the model measurements as related to the HF case. This is necessary since the angle of the lobes (for small departure angles) are approximately proportional to $\sqrt{1-\frac{75}{150}}$ where 75 is the length of the long wire radials and 150 is the horizontal separation between transmitting antenna and the receiving monopole (see fig. 1). If the separation between the transmitting antenna and receiving monopole increases from 150 m to infinity as in the field case,

the lobes are raised in elevation. For this experiment $\sqrt{1-\frac{75}{150}}=\frac{1}{2}$. Thus the experimentally measured lobe positions should be raised by more than 40 percent for optimum siting of a point-to-point HF link.

Another consideration to be made when interpreting the scaled model measurements as regards the HF case is the influence of the earth curvature. At these low angles the pattern is cut back due to diffraction. For the 10 Mc/s HF case this would amount to approximately ½ degree [Wait and Conda, 1958b]. Thus some error could be expected due to the earth curvature, particularly at low angles of the order of 1 degree.

5. Conclusions

The results of these measurements show that an increase in the density (number of wires in the sector) improves the relative level of the received signal.

Measurements of ground conductivity, effect of increased wire density, and effect of even longer wire radial ground systems were beyond the scope of these measurements. In order to completely investigate the advantages of ground systems with extended wire radials, it would be desirable to make a further study with an additional number of wires and wire lengths, and to carry out measurements at higher angles of arrival.

This work was carried out in 1959 in support of the HF Arctic Propagation Project No. 85421 sponsored by the U.S.N. BuShips.

The measurements were made with the assistance of W. L. Martin, G. W. Angus, and R. W. Harsh, of the National Bureau of Standards, and W. E. McKenney and F. A. Foss of Scatter Communication Corporation.

6. References

Laport, E. A., Radio Antenna Eng., pp. 82 and 117 (McGraw-Hill Book Co. Inc., New York, N.Y., 1952).

Wait, J. R. and A. M. Conda, The patterns of a slot-array on a finite ground plane, L'Onde Electrique Special Supplement, Vol. 1, pp. 21–29, Proc. of the International Congress on UHF Circuits and Antennas, Paris (1958a).

Wait, J. R., and A. M. Conda, Pattern of an antenna on a curved lossy surface, IRE Trans. on Antennas and Prop. **AP–6,** No. 4, 348–359 (Oct. 1958b).

7. Appendix[1]

7.1. Path Geometry

A very important consideration for evaluating measurements as described in this paper is that the target transmitter is at a finite distance whereas the actual source of the field for a full size antenna is a downcoming plane wave (or at least something approaching it). Figure 4 shows the geometry of the path.

[1] This appendix is based on a private communication from J. R. Wait.

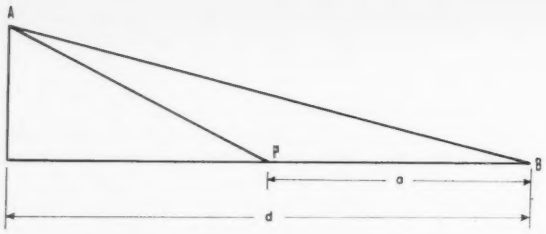

FIGURE 4. *Drawing of measurement path geometry.*
Transmitting antenna at point A; receiving monopole at point B.

On the assumption that PB is a well conducting ground plane it can be expected that the antenna pattern first maximum in the vertical plane occurs when $\Delta=AP+PB-AB=m^2\frac{\lambda}{4}$ when m^2 is some number of order unity.

By trigonometry the following will be obtained:

$$\Delta=[(d\tan\psi)^2+(d-a)^2]^{1/2}+a-\frac{d}{\cos\psi},$$

If ψ is small this simplifies to

$$\Delta\simeq\frac{a\psi^2}{2}\left[\frac{1}{1-\frac{a}{d}}\right],$$

The angle ψ which satisfies $\Delta=m\frac{\lambda}{4}$ is

$$\psi_0\simeq m\left\{\frac{\lambda}{2a}\left[1-\frac{a}{d}\right]\right\}^{1/2},$$

The value of m can be identified with the value of u where the function

$$F(u)=\frac{1+i}{2}\left\{\left[\frac{1}{2}+C(u)\right]-i\left[\frac{1}{2}+S(u)\right]\right\}.$$

is a maximum. C and S are Fresnel integrals [Wait and Conda, 1958a]. Thus $m\simeq1.2$.

For the measurements carried out at a frequency of 400 Mc/s the first maximum is obtained. $a=75$ m thus $\frac{2a}{\lambda}\simeq14$ and

$$\psi_0\simeq\frac{m}{14}\left(1-\frac{a}{d}\right)^{1/2}\text{ radians}$$

$$\simeq4m\left(1-\frac{a}{d}\right)^{1/2}\text{ degrees}$$

$$\psi_0\simeq2.8m\simeq3.3°$$

For the second maximum $m\simeq2.2$ which gives $\psi_0\simeq6.2°$. For the first null $m\simeq1.8$ which gives $\psi_0\simeq5.0°$.

7.2. Earth Curvature Influence

Another consideration in evaluating the described measurements as pertained to an actual operational antenna is the influence of the earth's curvature [Wait and Conda, 1958b].

At these low angles the pattern is "cut back" due to diffraction. The general restriction on neglecting this effect is then

$$\psi >> \left(\frac{2}{ka}\right)^{1/3} \text{ radians}$$

where ka is the circumference of the earth in wavelengths. For 10 Mc/s this works out to be $\psi \gg \frac{1}{2}$ degree. Thus some error could be expected, particularly at low angles of the order of 1 degree.

(Paper 65D6–167)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 6, November–December 1961

Publications of the Staff of the National Bureau of Standards*

Selected Abstracts

Some boundary value problems involving plasma media, J. R. Wait, *J. Research NBS* **65B** (*Math. and Math. Phys.*) *No. 2, 137–150 (April–June 1961).*
A plasma, consisting of a neutral mixture of electrons, ions, and molecules, in the presence of a constant magnetic field H_0, possesses a dielectric constant which is in the form of a tensor. Exact solutions of boundary value problems involving such media are obtained for two-dimensional configurations. Explicit results are given for the reflection coefficients of stratified plasma in planar and cylindrical geometry.

A survey of spread-F, F. N. Glover, *NBS TN 82 (PB 161583) (Nov. 1960)* $1.75.
Examples of spread-*F* forms occurring at different latitudes are presented, illustrating the classification of spread into range type and frequency type. The occurrence patterns of spread-*F* at different latitudes are correlated with other geophysical phenomena. Magnetic latitude and time within the sunspot cycle appreciably affect the pattern of spread occurrence. Instrumental techniques and their advantages for spread studies are outlined. The principal theoretical explanations of spread-*F* are summarized. A single mechanism need not be postulated as responsible for all types of spread occurrence or at all latitudes.

A diffraction theory for LF sky-wave propagation, J. R. Wait, *J. Geophys. Research* **66,** *No. 6, 1713–1724 (June 1961).*
The concept that radio waves propagate from transmitter to receiver via discrete ionospheric reflections (i.e., hops) is put on a sound theoretical basis. The earth and the ionosphere are represented by two (smooth) concentric spherical surfaces. The (local) reflecting characteristics of these boundaries are assumed to be known. It is shown that the *m*th hop sky wave may be generally represented by a complex integral. The geometrical-optical representations for the field are retrieved as a special case. For low radio frequencies, it is shown that these optical-type formulas, which are often used, have severe limitations. Generally, they are not valid near the caustic points. Techniques for evaluating the complex integral for the sky waves near and beyond the caustics are then discussed in some detail. Where possible, a physical interpretation of the results is given.

A diffraction theory for LF sky-wave propagation—an additional note, *J. R. Wait and A. M. Conda, J. Geophys. Research* **66,** *No. 6, 1725–1729 (June 1961).*
The behavior of the first-hop sky wave near its geometrical horizon is discussed. Using formulas already developed, numerical results in this difficult caustic region are presented in graphical form. It is shown that, sufficiently far away from the caustic, the field approaches its geometrical-optical value. The results should be useful for sky-wave field computation at low frequencies and in the interpretation of measured field-strength data.

Photoionization heating in the F region of the atmosphere, *D. C. Hunt and T. E. Van Zandt, J. Geophys. Research* **66,** *No. 6, 1673–1682 (June 1961).*
The thermal economy in the *F* region of the atmosphere is investigated using the barometric equation and the continuity equation for thermal energy. It is concluded that photoionization is certainly a major source, and possibly the predominant source, of *F*-region heating. The assumptions and approximations are discussed and are shown not to invalidate the conclusion. The heat flux due to photoionization in the *F* region at low and middle latitudes at sunspot maximum is estimated to lie between 0.5 and 2.0 ergs/cm² sec, with a most probable value of about 1.0 erg/cm² sec.

The use of interferometer observations of satellites for measurement of irregular ionospheric refraction, R. S. Lawrence and J. W. Warwick, *Annals of the Intern. Geophys. Year* **12,** Pt. 2, pp. 566–569 (Pergamon Press, New York, N.Y., 1961).
We outline a digital method for determination of satellite orbital parameters from a single passage through the lobes of a simple interferometer. Differences between the observed and computed angle-of-arrival indicate that, at least for low-altitude passages, the ionospheric irregularities probably occur in a thin layer and are not strongly aligned along the magnetic field.

Long-distance one-hop F_1 propagation through the auroral zone, L. H. Tveten, *J. Geophys. Research* **66,** *No. 6, 1683–1684* (June 1961).
The existence of one-hop F_1 propagation through the auroral zone over a 4,495-km path is described, and an example of a sweep-frequency record illustrating this F_1 mode is shown.

Incoherent scattering by free electrons as a technique for studying the ionosphere and exosphere: Some observations and theoretical considerations, K. L. Bowles, *pp. 223–242 of AGARDograph* **42, The upper atmosphere above F2-maximum,** *North Atlantic Treaty Organization, Advisory Group for Aeronautical Research and Development, Avionics Panel. Papers presented at the Symp. Ionospheric Research Committee, Paris, France, May 1959. (Technical Editing and Reproduction Ltd., 95 Great Portland Street, London, W. 1, England.)*
Incoherent scattering by the free electrons of the ionosphere has been suggested as a technique for measuring the electron density profile both below and above the *F*-region maximum. This paper reports observations which confirm the existence of the incoherent scatter and show that its intensity is essentially the predicted value. The observed Doppler broadening is considerably smaller than predicted. In the second part of the paper, an explanation for the reduced Doppler broadening is offered. The scatter is explained as arising from statistical fluctuations of electron density, the distribution of which is controlled by the positive ions.

A model of the *F*-region above $h_{max}F_2$, J. W. Wright, *pp. 211–222 of AGARDograph* **42, The upper atmosphere above F2-maximum,** *North Atlantic Treaty Organization, Advisory Group for Aeronautical Research and Development, Avionics Panel. Papers presented at the Symp. Ionospheric Research Committee, Paris, France, May 1959.*
A simple Chapman model (scale height 100 km) of the ionosphere above the peak of the *F*-region is found to be in good agreement with the few observed profiles of this region. The ratio of the above-peak electron content implied by the model, to below-peak electron content, is found to vary from about 2.8 to 4.0, also in good agreement with observations. The model is used in conjunction with electron density profiles observed from the ground to construct meridional cross-sections along the 75° W geographic meridian.

A comparison between theoretical and experimental data on phase velocity of VLF radio waves, J. R. Wait, *Proc. IRE* **49,** *No. 6, 1089–1090 (June 1961).*
It can be concluded that there is fairly good agreement between measured phase velocity and the mode theory as developed for a curved earth-ionosphere waveguide. Furthermore, the heights of the equivalent reflecting layer are fairly consistent with other results.

Other NBS Publications

Journal of Research NBS 65A (Phys. and Chem.) No. 5 (September–October 1961) 75 cents.

Calibration of a monitor for use in bremsstrahlung beams. E. G. Fuller and E. Hayward.
Mass spectrometric study of NF_2, NF_3, N_2F_2, and N_2F_4. J. T. Herron and V. H. Dibeler.
Rate of the reaction NO+N, and some heterogeneous reactions observed in the ion source of a mass spectrometer. J. T. Herron.
Synthesis of the humites $nMg_2SiO_4 \cdot Mg(F,OH)_2$. A. Van Valkenburg.
Phase equilibria in systems involving the rare earth oxides. Part III. The Eu_2O_3-In_2O_3 system. S. J. Schneider.
Heats of hydrolysis and formation of dimethoxychloroborane. M. V. Kilday, W. H. Johnson, and E. J. Prosen.
Tritium-labeled compounds VII. Isotope effects in the oxidation of D-mannitols-C^{14} and D-mannitols-*t* to D-fructoses. L. T. Sniegoski, H. L. Frush, and H. S. Isbell.
Franck-Condon factors to high vibrational quantum numbers I: N_2 and N_2^+. R. W. Nicholls.

Journal of Research NBS 65B (Math. and Math. Phys.) No. 3 (July–September 1961) 75 cents.

Theory of an accurate intermediary orbit for satellite astronomy. J. P. Vinti.
Note on the "baffled piston" problem. F. Oberhettinger.
Some results on non-negative matrices. M. Marcus, H. Minc, and B. N. Moyls.
Probability inequalities of the Tchebycheff type. I. R. Savage.

Journal of Research, NBS 65C (Eng. and Instr.) No. 3 (July–September 1961) 75 cents.

Prediction of symptoms of cavitation. R. B. Jacobs.
Heating and cooling of air flowing through an underground tunnel. B. A. Peavy.
Stress-corrosion cracking of the AZ31B magnesium alloy. H. L. Logan.
Coatings formed on steel by cathodic protection and their evaluation by polarization measurements. W. J. Schwerdtfeger and R. J. Manuele.
Calibration of inductance standards in the Maxwell-Wien bridge circuit. T. L. Zapf.
Calibration of loop antennas at VLF. A. G. Jean, H. E. Taggart, and J. R. Wait.
Location of the plane of best average definition with low contrast resolution patterns. F. E. Washer and W. P. Tayman.
Influence of temperature and relative humidity on the photographic response to Co^{60} gamma radiation. M. Ehrlich.

Ideal gas thermodynamic functions and isotope exchange functions for the diatomic hydrides, deuterides, and tritrides, L. Haar, A. S. Friedman, and C. W. Beckett, NBS Mono. 20 (1961) $2.75.
Bibliography of temperature measurement—January 1953 to June 1960, C. Halpern and R. J. Moffat, NBS Mono. 27 (1961) 15 cents.
Causes of variation in chemical analyses and physical tests of portland cement, B. L. Bean and J. R. Dise, NBS Mono. 28 (1961) 25 cents.
Thermal expansion of technical solids at low temperatures. A compilation from the literature, R. J. Corruccini and J. J. Gniewek, NBS Mono. 29 (1961) 20 cents.
Corrected optical pyrometer readings, D. E. Poland, J. W. Green, and J. L. Margrave, NBS Mono. 30 (1961) 55 cents.
Publications of the National Bureau of Standards July 1, 1957, to June 30, 1960 (Includes Titles of Papers Published in Outside Journals 1950 to 1959), B. L. Arnold, NBS Misc. Publ. 240 (1961) $2.25.
Mean electron density variations of the quiet ionosphere—May 3, 1959, J. W. Wright, L. R. Wescott and D. J. Brown, NBS TN40–3 (PB151399–3) (1960) $1.50.
A survey of computer programs for chemical information searching, E. C. Marden and H. R. Koller, NBS TN85 (PB161586) (1961) $2.25.
Integrated starlight over the sky F. E. Roach, and L. R. Megill. Astrophys. J. **133,** No. 1, 228–242 (Jan. 1961).
The electromagnetic fields of a dipole in the presence of a thin plasma sheet, J. R. Wait, Appl. Sci. Research **8** sec. B, 397–417 (1960).
The nature, cause and effect of the porosity in electrodeposits, VII. A microscopic examination of nickel-chromium coatings after atmospheric corrosion, D. W. Ernst and F. Ogburn, Plating **48,** No. 4, 491–497 (May 1961).
Physical quantities proposed for radiation measurements, L. S. Taylor, Proc. IXth Intern. Congress of Radiology, July 23–30, 1959 (Munchen, Germany), 1301–1302 (1960).
Improving rectifier circuits, G. F. Montgomery, Electronics p. 86–87 (April 7, 1961).
Fracture characteristics of notched tensile specimens of titanium and a titanium alloy, G. W. Geil and N. L. Carwile, Materials Research and Standards (Am. Soc. Testing Materials, Philadelphia, Pa.), **1,** No. 1, 16–21 (Jan. 1961).
The robustness of life testing procedures derived from the exponential distribution, M. Zelen and M. C. Dannemiller, Technometrics **3,** No. 1, 29–49 (Feb. 1961).
Absence of an isotope effect in the fractional recrystallization of alpha-*D*-glucose-1-*t*, H. S. Isbell, H. L. Frush, and N. B. Holt, Anal, Chem. **33,** 225–226 (Feb. 1961).
Applications of graphs and Boolean matrices to computer programming, R. B. Marimont, SIAM Rev. **2,** No. 4, 259–268 (Oct. 1960).
Comparison of national standards for roentgen measurement, H. O. Wyckoff, IXth Intern. Congress of Radiology, July 23–30, 1959 (Munchen, Germany), 1315–1318 (1960).
Environmental factors in a family-size underground shelter, R. P. Achenbach, Proc. Meeting on Environmental Engineering in Protective Shelters, Feb. 8–10, 1960, Natl. Acad. Sci.-Natl. Research Council, Wash. D.C., 69–118 (1961).
How to evaluate accuracy, W. J. Youden, Materials Research and Standards (Am. Soc. Testing Materials, Philadelphia, Pa.), **1,** No. 4, 268–271 (April 1961).
Improved standard for the calibration of vibration pickups, R. R. Bouche, Exp. Mech. 1–6 (April 1961).
Absorption and scattering of photons by holmium and erbium, E. G. Fuller and E. Hayward, Proc. Intern. Conf. Nuclear Structure, 763–766 (Kingston, Canada, 1960).
Introductory remarks, U. Fano, Conf. Coherence Properties of Electromagnetic Radiation, AFOSR–583, Report on Conf., 31–38 (University of Rochester, The Institute of Optics, Rochester, N.Y., April 1961).
Available heat sinks for protected underground installations, B. A. Peavy, Proc. Meeting on Environmental Engineering in Protective Shelters, Feb. 8–10, 1960, Natl. Acad. Sci.-Natl. Research Council, Wash. D.C., 69–118 (1961).
Interpretation of the appearance potentials of secondary ions, M. B. Wallenstein and M. Krauss, J. Chem. Phys. **34,** No. 3, 929–936 (Mar. 1961).
Observations on the chemiluminescent reaction of nitric oxide with atomic oxygen, H. P. Broida, H. I. Schiff, and T. M. Sugden, Trans. Faraday Soc. **57,** No. 458, 259–265 (Feb. 1961).
Studies of the low-temperature distillation of hydrogen isotopes, T. M. Flynn, Cryogenics **1,** No. 2, 1–5 (Dec. 1960); Advances Cryogenic Engr. **6,** 236–244 (1961).
Inert enclosed pump for shaped flow of ultraclean solutions, M. W. Barnes and R. H. Noyce, Rev. Sci. Instr. **32,** No. 3, 353 (Mar. 1961).
On some partial differential equations of Brownian motion of a free particle, A. Ghaffari, (Abstract) Proc. Intern. Conf. on Partial Differential Equations and Continuum Mechanics (Univ. of Wisconsin, Madison, Wisc., June 7–15, 1960), 348–350 (1961).
Radiation protection standards, L. S. Taylor, Radiology **74,** 824–831 (May 1960).
Atomic beam frequency standards, R. C. Mockler, R. E. Beehler, and C. S. Snider, IRE Trans. Instrumentation **I–9,** No. 2, 120–132 (Sept. 1960).
The development of more stable gage blocks, M. R. Meyerson, T. R. Young, and W. R. Ney, ASTM Bull. **1,** No. 5, 368–374 (May 1961).

The power spectrum and its importance in precise frequency measurements, J. A. Barnes and R. C. Mockler, IRE Trans. Instrumentation **I-9,** No. 2, 149–155 (Sept. 1960).

VLF phase perturbation associated with meteor shower ionization, C. J. Chilton, J. Geophys. Research **66,** No. 2, 379–383 (Feb. 1961).

Departures from the Saha equation for ionized helium. I. Condition of detailed balance in the resonance lines, R. N. Thomas and J. B. Zirker, Astrophys. J. **133,** No. 2, 588–595 (Mar. 1961).

Tests for regression coefficients when errors are correlated, M. M. Siddiqui, Annals Math. Stat. **31,** No. 4, 929–938 (Dec. 1960).

Electron diffraction studies on solid α-nitrogen, E. M. Horl and L. Marton, Acta Cryst. **14,** Pt. I, 11–19 (Jan. 10, 1961).

A comment on the NRL solar Lyman-Alpha results, J. T. Jeffries and R. N. Thomas, Astrophys. J. **133,** No. 2, 606–607 (Mar. 1961).

The nation's electronic standards program: Where do we now stand? H. W. Lance, IRE Trans. Instrumentation **I-9,** No. 2, 94–100 (Sept. 1960).

Hard gallium alloys for use as low contact resistance electrodes and for bonding thermocouples into samples, G. C. Harman, Rev. Sci. Instr. **31,** No. 7, 717–720 (July 1960).

Scattering of photons by deformed nuclei, E. G. Fuller and E. Hayward, Proc. Intern, Conf. Nuclear Structure, 763–766 (Kingston, Canada, 1960).

Other CRPL Publications

Evidence on the laminar nature of the exosphere obtained by means of guided high-frequency wave propagation, R. M. Galley and W. F. Utlaut, Phys. Rev. Letters **6,** No. 11, 591–594 (June 1, 1961).

Solar flare effects in the *F* region of the ionosphere (letter to the editor) R. W. Knecht and K. Davies, Nature **190,** No. 4778, 797–798 (May 27, 1961).

Summary of outstanding solar events in July 1959, A. H. Shapley and Dorothy Trotter, Union Géodésique et Géophysique Internationale, Symposium on the July 1959 Events and Associated Phenomena, Helsinki, July 1960 (L'Institut Geographique National, Paris, 1961), pp. 72–76.

Solar disturbances and radio communication forecasts James F. Brockman, Sky and Telescope **XXI,** No. 6, 322–326 (June 1961).

A note on the phase velocity of radio waves, J. R. Wait and K. P. Spies, J. Geophys. Research **66,** 992–993 (March 1961).

**Publications for which a price is indicated (except for Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (order by PB number). Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

INDEX TO VOLUME 65D

January to December 1961

A

Page

A few observations of the perturbations in the phase of the low-frequency ground wave ... 393
A quick method for estimating the stage of the sunspot cycle ... 637
A new approach to the mode theory of VLF propagation ... 37
A note concerning the excitation of ELF electromagnetic waves ... 481
Abraham, L. G., Jr., J. A. Bradshaw, Weather and reception level on a troposphere link—annual and short-term correlations ... 155
Absorption, radio wave, of several gases in the 100 to 117 kMc/s frequency range ... 15
Agy, V., K. Davies, R. Salaman, An atlas of oblique-incidence ionograms (a digest) ... 35
Airglow excitation by gyrowaves, power density requirements ... 321
Almost fifty years of URSI ... 317
Amplitude and angular scintillations of the radio source Cygnus-A observed at Boulder, Colo ... 333
Amplitude distribution of VLF atmospheric noise, effect of receiver bandwidth ... 299
Analysis of LF ionospheric radio propagation phenomena ... 507
Angles upon HF radio signals propagated over long distances, effect of antenna radiation ... 167
Angular and amplitude scintillations of the radio source Cygnus-A observed at Boulder, Colo ... 333
Antenna coupling error in direction finders ... 363
Antenna, high-resolution rapid-scan ... 101
Antenna radiation angles upon HF radio signals propagated over long distances, effect ... 167
Antenna, underground, useful radiation ... 89
Antenna with a circular conducting-disk ground system on the surface of a lossy half-space, impedance of a monopole ... 183
Antennas, circular aperture, Fresnel region fields ... 137
Appleton-Hartree formula, validity of some approximations ... 323
Approximations to the Appleton-Hartree formula, validity ... 323
Atlas of oblique-incidence ionograms (a digest) ... 35
Atmosphere, formula for radio ray refraction in an exponential ... 181
Atmosphere, graphical determination of radio ray bending in an exponential ... 175
Atmospheric radio noise predictions, reliability ... 565
Attenuation coefficients for propagation at very low frequencies (VLF) during a sudden ionospheric disturbance (SID) ... 543
Attenuation rate of ELF radio waves, observed ... 475
Auroral arcs, red, and ionospheric recombination, relationship ... 129
Automatic direction finding, design for spinning goniometer ... 237

B

Bailey, V. A., Power density requirements for airglow excitation by gyrowaves ... 321
Bain, W. C., Phase difference observations at spaced aerials and their application to direction-finding ... 229
Barsis, A. P., R. S. Kirby, VHF and UHF signal characteristics observed on a long knife-edge diffraction path ... 437
Beale, E. M. L., Brooke variance classification system for DF bearings ... 255
—, Estimation of variances of position lines from fixes with unknown target positions ... 263
Bean, B. R., B. A. Cahoon, Correlation of monthly median transmission loss and refractive index profile characteristics ... 67
Blackband, W. T., Effects of the ionosphere on VLF navigational aids ... 575
Blair, J. C., R. N. Davis, Jr., R. C. Kirby, Frequency dependence of *D*-region scattering at VHF ... 417
Blair, W. E., R. K. Moore, Dipole radiation in a conducting half space ... 547
Boules, K. L., Incoherent scattering by free electrons as a technique for studying the ionosphere and exosphere: some observations and theoretical considerations ... 1
Bradshaw, J. A., L. G. Abraham, Jr., Weather and reception level on a troposphere link—annual and short-term correlations ... 155
Britt, C. O., C. W. Tolbert, A. W. Straiton, Radio wave absorption of several gases in the 100 to 117 kMc/s frequency range ... 15
Brooke variance classification system for DF bearings ... 255
Bowhill, S. A., Diversity effects in long distance high frequency radio pulse propagation ... 213
—, Statistics of a radio wave diffracted by a random ionosphere ... 275
Brown, W. W., A. D. Watt, R. W. Plush, A. H. Morgan, Worldwide VLF standard frequensy and time signal broadcasting ... 617

C

Cahoon, B. A., B. R. Bean, Correlation of monthly median loss and refractive index profile characteristics ... 67
Chadwick, W. B., A quick method for estimating the stage of the sunspot cycle ... 637
Characteristics of waveguides for long-distance transmission ... 75
Classification system for DF bearings, Brooke variance ... 255
Collision frequencies of free electrons in an ionized region, the electrically short antenna for measuring ... 371
Comparison between mode theory and ray theory of VLF propagation ... 357
Composite cylinder, on the theory of diffraction ... 19
Computation of whistler ray paths ... 485
Conducting half space, dipole radiation ... 547
Correlation arrays, resolution characteristics ... 245
Corelation of monthly median transmission loss and refractive index profile characteristics ... 67
Correlations, annual and short-term, weather and reception level on a troposphere link ... 155
Cottony, H. V., A. C. Wilson, A high-resulution rapid-scan antenna ... 101
Coupled variables, nonsingular embedding of transition processes within a more general framework ... 595
Crombie, D. D., Reflection from a sharply bounded ionosphere for VLF propagation perpendicular to the magnetic meridian ... 455
Cygnus-A observed at Boulder, Colo., amplitude and angular scintillations of the radio source ... 333
Cylinder, on the theory of diffraction by a composite ... 19

D

Davies, K., King, G. A. M., On the validity of some approximations to the Appleton-Hartree formula ... 323
—, *R. Salaman, V. Agy*, An atlas of oblique-incidence ionograms (a digest) ... 35
Davis, R. N., Jr., J. C. Blair, R. C. Kirby, Frequency dependence of *D*-region scattering at VHF ... 417
Decker, M. T., W. J. Hartman, Mutual interference between surface and satellite communication systems ... 433
Dellinger, J. H., Almost fifty years of URSI ... 317
Denton, D. H., R. W. P. King, C. W. Harrison, The electrically short antenna as a probe for measuring free electron densities and collision frequences in an ionized region ... 371
Design for spinning goniometer automatic direction finding ... 237
Design of panoramic ionospheric resources ... 629
DF bearings, Brooke variance classification system ... 255
Diffraction by a composite cylinder ... 19
Digital methods for the extraction of phase and amplitude information from a modulated signal ... 351
Dipole, excitation of VLF and ELF radio waves by a horizontal magnetic ... 305
Dipole radiation in a conducting half space ... 547
Direction finders, antenna coupling error ... 363
Direction-finding, phase difference observations at spaced aerials and their application ... 229
Direction-finding problems, long-distance location ... 233
Direction-finding techniques; propagation studies ... 197
Distance high frequency radio pulse propagation, diversity effects ... 213
Diversity effects in long distance high frequency radio pulse propagation ... 213
Dobrott, D., A. Ishimaru, East-west effect on VLF mode transmission across the earth's magnetic field ... 47
Doherty, R. H., G. Hefley, R. F. Linfield, Initial results of a new technique for investigating sferic activity ... 157
D-region scattering at VHF, frequency dependence of ... 417

E

Earth and ionosphere, resonance of the space ... 465
East-west effect on VLF mode transmission across the earth's magnetic field ... 47
Effect of antenna radiation angles upon HF radio signal propagated over long distances ... 167
Effect of multiple atmospheric inversions on tropospheric radio propagation ... 385
Effect of receiver bandwidth on the amplitude distribution of VLF atmospheric noise ... 299
Effects of the ionosphere on VLF navigational aids ... 575
Electrically short antenna as a probe for measuring free electron densities and collision frequencies in an ionized region ... 371
Electromagnetic waves, note concerning the excitation of ELF ... 481
Electron densities, electrically short antenna for measuring ... 371
ELF and VLF radio waves by a horizontal magnetic dipole, excitation ... 305
ELF electromagnetic waves, a note concerning the excitation ... 481
ELF radio waves, observed attentuation rate ... 475
Embedding of transition processes, nonsingular, within a more general framework of coupled variables ... 595
Error in direction finders, antenna coupling ... 363
Estimation of variances of position lines from fixes with unknown target positions ... 263
Excitation of ELF electromagnetic waves, a note concerning ... 481
Excitation of VLF and ELF radio waves by a horizontal magnetic dipole ... 305
Exosphere and ionosphere, incoherent scattering by free electrons as a technique for studying: some observations and theoretical considerations ... 1
Experimental study of inverted L-, T-, and related transmission-line antennas ... 449
Extraction of phase and amplitude information from a modulated signal, digital methods ... 351
Extremely low frequencies, on the spectrum of terrestrial radio noise ... 581

F

F-layer and sporadic-*E* at VHF in the Far East, observation ... 93
Formula for radio ray refraction in an exponential atmosphere ... 181
Free-balloon borne meterological refractometer ... 149
Frequency dependence of *D*-region scattering at VHF ... 417
Fresnel region fields of circular aperture antennas ... 137
Fulton, F. F., Jr., Effect of receiver bandwidth on the amplitude distribution of VLF atmospheric noise ... 299

G

Galejs, J., Excitation of VLF and ELF radio waves by a horizontal magnetic dipole ... 305
Gases, radio wave absorption, in the 100 to 117 kMc/s frequency range ... 15
Gething, P. J. D., Influence of ionospheric conditions on the accuracy of high frequency direction finding ... 225
Gossard, E. E., J. F. Theisen, Free-balloon borne meteorological refractometer ... 149
Graphical determination of radio ray bending in an exponential atmosphere ... 175
Ground wave, low-frequency phase, a few observations of the perturbations ... 393
Ground-wave propagation on a spherical earth, theory of mixed-path ... 401

H

Page

Harrison, C. W. Jr., Antenna coupling error in direction finders ... 363
—, *R. W. P. King, D. H. Denton,* The electrically short antenna as a probe for measuring free electron densities and collision frequencies in an ionized region ... 371
Hartman, W. J., M. T. Decker, Mutual interference between surface and satellite communication systems ... 433
Hayden, E. C., Instrumentation for propagation and direction-finding measurements ... 253
—, Propagation studies using direction-finding techniques ... 197
Hayre, H. S., R. K. Moore, Theoretical scattering coefficient for near vertical incidence from contour maps ... 427
Heading, J., Nonsingular embedding of transition processes within a more general framework of coupled variables ... 595
Hefley, G., R. H. Doherty, R. F. Linfield, Initial results of a new technique for investigating sferic activity ... 157
Heim, D. S., W. J. Lindsay, Design for spinning goniometer automatic direction finding ... 237
Heisler, L. H., L. D. Wilson, Design of panoramic ionospheric resources ... 629
Herman, J. R., Reliability of atmospheric radio noise predictions ... 565
High frequency backscatter sounders, ionospheric motions ... 115
High frequency direction finding, influence of ionospheric conditions on the accuracy of ... 225
High frequency radio pulse propagation, diversity effects in long distance ... 213
High-resolution rapid-scan antenna ... 101
Horizontal polarization, smooth earth diffraction calculations ... 397
Hu, Ming-Kuei, Fresnel region fields of circular aperture antennas ... 137

I

Impedance of a monopole antenna with a circular conducting-disk ground system on the surface of a lossy half-space ... 183
Incoherent scattering by free electrons as a technique for studying the ionosphere and exosphere: some observations and theoretical considerations ... 1
Influence of ionospheric conditions on the accuracy of high frequency direction finding ... 225
Initial results of a new technique for investigating sferic activity ... 157
Instrumentation for propagation and direction-finding measurements ... 253
Interference, mutual, between surface and satellite communication systems ... 433
Inverted L-, T-, and related transmission-line antennas, experimental study ... 449
Ionograms (a digest), an atlas of oblique-incidence ... 35
Ionosphere and earth, resonance of the space ... 465
Ionosphere and exosphere, incoherent scattering by free electrons as a technique for studying: some observations and theoretical considerations ... 1
Ionosphere, magneto-ionic propagation phenomena in low- and very-low-radiofrequency waves reflected ... 53
Ionospheric conditions on the accuracy of high frequency direction finding, influence ... 225
Ionospheric motions observed with high frequency backscatter sounders ... 115
Ionospheric recombination and red auroral arcs, relationship ... 129
Ishikawa, M., K. Miya, T. Sasaki, Observations of *F*- layer and sporadic-*E* scatter at VHF in the Far East ... 93
Ishimaru, A., D. Dobrott, East-west effect on VLF mode transmission across the earth's magnetic field ... 47

J

Jean, A. G., A. C. Murphy, J. R. Wait, D. F. Wasmundt, Observed attenuation rate of ELF radio waves ... 475
Jespersen, J. L., R. S. Lawrence, R. C. Lamb, Amplitude and angular scintillations of the radio source Cygnus-A observed at Boulder, Colo ... 333
—, —, —, Digital methods for the extraction of phase and amplitude information from a modulated signal ... 351
Johler, J. R., Magneto-ionic propagation phenomena in low- and very-low-radiofrequency waves reflected by the ionosphere ... 53
—, On the analysis of LF ionospheric radio propagation phenomena ... 507

K

Karbowiak, A. E., L. Solymar, Characteristics of waveguides for long-distance transmission ... 75
King, G. A. M., K. Davies, On the validity of some approximations to the Appleton-Hartree formula ... 323
—, *F. E. Roach,* Relationship between red auroral arcs and ionospheric recombination ... 129
King, R. J., S. W. Maley, Impedance of a monopole antenna with a circular conducting-disk ground system on the surface of a lossy half-space ... 183
King, R. W. P., C. W. Harrison, D. H. Denton, The electrically short antenna as a probe for measuring free electron densities and collision frequencies in an ionized region ... 371
—, *S. Prasad,* Experimental study of inverted L-, T-, and related transmission-line antennas ... 449
Kirby, R. C., J. C. Blair, R. N. Davis, Jr., Frequency dependence of *D*-region scattering at VHF ... 417
Kirby, R. S., A. P. Barsis, VHF and UHF signal characteristics observed on a long knife-edge diffraction path ... 437
Kirch, J. E., J. M. Foss, A few observations of the perturbations in the phase of the low-frequency ground wave ... 393
Kirchhoff-Huygens' principle ... 427
Knife-edge diffraction path, VHF and UHF signal characteristics observed ... 437
Kodis, R. D., On the theory of diffraction by a composite cylinder ... 19

L

Lamb, R. C., R. S. Lawrence, J. L. Jespersen, Amplitude and angular scintillations of the radio source Cygnus-A observed at Boulder, Colo ... 333
—, —, —, Digital methods for the extraction of phase and amplitude information from a modulated signal ... 351

Page

Lawrence, R. S., J. L. Jespersen, R. C. Lamb, Amplitude and angular scintillations of the radio source Cygnus-A observed at Boulder, Colo ... 333
—, —, —, Digital methods for the extraction of phase and amplitude information from a modulated signal ... 351
Linder, I. W., Resolution characteristics of correlation arrays ... 245
Lindsay, W. J., D. S. Heim, Design for spinning goniometer automatic direction finding ... 237
Lines from fixes with unknown target positions, estimation of variances ... 263
Linfield, R. F., G. Hefley, R. H. Doherty, Initial results of a new technique for investigating sferic activity ... 157
LF ionospheric radio propagation phenomena, analysis ... 507
Long-distance location and direction-finding problems ... 233
Low-angle radiation from a monopole, measurements ... 641
Low-frequency ground wave phase perturbations, a few observations ... 393

M

Magnetic field, Earth's east-west effect on VLF mode transmission ... 47
Magnetic meridian, reflection from a sharply bounded ionosphere for VLF perpendicular propagation ... 455
Magneto-ionic propagation phenomena in low- and very-low-radiofrequency waves reflected by the ionosphere ... 53
Maley, S. W., R. J. King, Impedance of a monopole antenna with a circular conducting-disk ground system on the surface of a lossy half-space ... 183
Measurements of low-angle radiation from a monopole ... 641
Measurements, propagation and direction-finding instrumentation ... 253
Meteorological refractometer, free-balloon borne ... 149
Mixed-path ground-wave propagation on a spherical earth ... 401
Miya, K., T. Sasaki, M. Ishikawa, Observation of *F*-layer and sporadic-*E* scatter at VHF in the Far East ... 93
Mode theory and ray theory of VLF propagation, comparison ... 357
Mode theory of VLF propagation, a new approach ... 37
Modulated signal, digital methods for the extraction of phase and amplitude information ... 351
Monopole antenna with a circular conducting-disk ground system on the surface of a lossy half-space, impedance ... 183
Moore, A. H., W. E. Blair, Dipole radiation in a conducting half space ... 547
—, *H. S. Hayre,* Theoretical scattering coefficient for near vertical incidence from contour maps ... 427
Morgan, A. H., A. D. Watt, R. W. Plush, W. W. Brown, Worldwide VLF standard frequency and time signal broadcasting ... 617
Motion characteristics of the backscatter ... 115
Multiple atmospheric inversions on tropospheric radio propagation, effect ... 385
Murphy, A. C., A. G. Jean, J. R. Wait, D. F. Wasmundt, Observed attenuation rate of ELF radio waves ... 475
Mutual interference between surface and satellite communication systems ... 433

N

Navigational aids, effects of the ionosphere on VLF ... 575
NBS research applicable to long-distance location and direction-finding problems ... 233
Near vertical incidence from contour maps, theoretical scattering coefficient ... 427
Nonsingular embedding of transition processes within a more general framework of coupled variables ... 595
Northover, F. H., Effect of multiple atmospheric inversions on tropospheric radio propagation ... 385

O

Oblique-incidence ionograms (a digest), an atlas ... 35
Observed attenuation rate of ELF radio waves ... 475
Observation of *F*-layer and sporadic-*E* scatter at VHF in the Far East ... 93
On the spectrum of terrestrial radio noise at extremely low frequencies ... 581
On the theory of diffraction by a composite cylinder ... 19
On the theory of mixed-path ground-wave propagation on a spherical earth ... 401
On the validity of some approximations to the Appleton-Hartree formula ... 323

P

Panoramic ionospheric resources, design ... 629
Pappas, C. F., L. E. Vogler, P. L. Rice, Graphical determination of radio ray bending in an exponential atmosphere ... 175
Parker, E. N., Solar wind ... 537
Path, VHF and UHF signal characteristics observed on a long knife-edge diffraction path ... 437
Paths, whistler ray, computation ... 485
Perturbations in the phase of the low-frequency ground wave, a few observations ... 393
Pierce, E. T., Attenuation of coefficients for propagation at very low frequencies (VLF) during a sudden ionosphere disturbance (SID) ... 543
Plush, R. W., A. D. Watt, W. W. Brown, A. H. Morgan, Worldwide VLF standard frequency and time signal broadcasting ... 617
Poeverlein, H., Resonance of the space between earth and ionosphere ... 465
Position lines from fixes with unknown target positions, estimation of variances ... 263
Power density requirements for airglow excitation by gyrowaves ... 321
Prasad, S., R. W. P. King, Experimental study of inverted L-, T-, and related transmission-line antennas ... 449
Propagation, a new approach to the mode theory of VLF ... 37
Propagation and direction finding measurements, instrumentation ... 253
Propagation at very low frequencies (VLF) during a sudden ionospheric disturbance (SID), attenuation coefficients ... 543
Propagation, diversity effects in long distance high frequency radio pulse ... 213
Propagation in the earth's crust, radio-wave ... 189
Propagation on a spherical earth, theory of mixed-path ground-wave ... 401
Propagation phenomena, magneto-ionic, in low- and very-low-radiofrequency waves reflected by the ionosphere ... 53
Propagation studies using direction-finding techniques ... 197

R

Page

Radar detection of surface roughness ... 427
Radiation, useful, from an underground antenna ... 89
Radio noise predictions, atmospheric, reliability ... 565
Radio propagation phenomena, on the analysis of I F ionospheric ... 507
Radio pulse propagation, diversity effects in long distance high frequency ... 213
Radio ray bending in an exponential atmosphere, graphical determination ... 175
Radio ray refraction in an exponential atmosphere, formula ... 181
Radio signals propagated over long distances, effect of antenna radiation angles upon HF ... 167
Radio signals, space analysis ... 293
Radio source Cygnus-A observed at Boulder, Colo., amplitude and angular scintillations ... 333
Radio wave absorption of several gases in the 100 to 117 kMc/s frequency range ... 15
Radio wave diffracted by a random ionosphere, statistics ... 275
Radio-wave propagation in the earth's crust ... 189
Radio waves by a horizontal magnetic dipole, excitation of VLF and ELF ... 305
Radio waves, observed attentuation rate of ELF ... 475
Radiofrequency waves, low- and very-low-, reflected by the ionosphere, magneto-ionic propagation phenomena ... 53
Raemer, H. R., On the spectrum of terrestrial radio noise at extremely low frequencies ... 581
Rapid-scan, high-resolution antenna ... 101
Reflection from a sharply bounded ionosphere for VLF propagation perpendicular to the magnetic meridian ... 455
Refraction in an exponential atmosphere, formula for radio ray ... 181
Refractive index profile characteristics and monthly median transmission loss, correlation ... 67
Refractometer, free-balloon borne meteorological ... 149
Related transmission-line antennas, experimental study of inverted L, T- ... 449
Relationship between red auroral arcs and ionospheric recombination ... 129
Reliability of atmospheric radio noise predictions ... 565
Research at NBS applicable to long-distance location and direction-finding problems ... 233
Resolution characteristics of correlation arrays ... 245
Resonance of the space between earth and ionosphere ... 465
Rice, P. L., C. F. Pappas, L. E. Vogler, Graphical determination of radio ray bending in an exponential atmosphere ... 175
Roach, F. E., G. A. M. King, Relationship between red auroral arcs and ionospheric recombination ... 129
Ross, J. M., J. E. Kirch, A few observations of the perturbations in the phase of the low-frequency ground wave ... 393

S

Salaman, R., V. Agy, K. Davies, An atlas of oblique-incidence ionograms (a digest) ... 35
Sasaki, T., K. Miya, M. Ishikawa, Observation of *F*-layer and sporadic-*E* scatter at VHF in the Far East ... 93
Satellite and surface communication systems, mutual interference ... 433
Scattering at VHF, frequency dependence of *D*-region ... 417
Scattering, incoherent, by free electrons as a technique for studying the ionosphere and exosphere: Some observations and theoretical considerations ... 1
Scintillations of the radio source Cygnus-A observed, amplitude and angular, at Boulder, Colo ... 333
Sferic activity, initial results of a new technique for investigating ... 157
Signal, VHF and UHF, characteristics observed on a long knife-edge diffraction path ... 437
Silberstein, R., Research at NBS applicable to long-distance location and direction-finding problems ... 233
Smooth earth diffraction calculations for horizontal polarization ... 397
Smyth, J. B., Space analysis of radio signals ... 293
Solar wind ... 537
Solymar, L. A. E. Karbowiak, Characteristics of waveguides for long-distance transmission ... 75
Space analysis of radio signals ... 293
Space between earth and ionosphere, resonance ... 465
Spaced aerials, and their application to direction-finding phase difference observations ... 229
Spherical earth, theory of mixed-path ground-wave propagation ... 401
Spinning goniometer automatic direction finding, design ... 237
Sporadic-*E* and *F*-layer scatter at VHF in the Far East, observation ... 93
Statistics of a radio wave diffracted by a random ionosphere ... 275
Straiton, A. W., C. O. Britt, C. W. Tolbert, Radio wave absorption of several gases in the 100 to 117 kMc/s frequency range ... 15
Sunspot cycle, a quick method for estimating the stage ... 637
Surface and satellite communication systems, mutual interference ... 433
Systems, mutual interference between surface and satellite communication ... 433

T

Target positions, estimation of variances of position lines from fixes with unknown ... 263
Terrain elevation autocovariance ... 427
Terrestrial radio noise at extremely low frequencies ... 581
Thayer, G. D., Formula for radio ray refraction in an exponential atmosphere ... 181
Theisen, J. F., E. E. Gossard, Free-balloon borne meteorological refractometer ... 149
Theoretical scattering coefficient for near vertical incidence from contour maps ... 427
Theory of diffraction by a composite cylinder ... 19
Tolbert, C. W., C. O. Britt, A. W. Straiton, Radio wave absorption of several gases in the 100 to 117 kMc/s frequency range ... 15
Transition processes, nonsingular embedding, within a more general framework of coupled variables ... 595
Transmission, long-distance, characteristics of waveguides ... 75
Transmission loss, monthly median, and refractive index profile characteristics, correlation ... 67
Transmission, VLF mode, east-west effect on across the earth's magnetic field ... 47
Troposphere link, weather and reception level—annual and short-term correlations ... 155
Tropospheric radio propagation, effect of multiple atmospheric inversions ... 385
Tveten, L. H., Ionospheric motions observed with high frequency backscatter sounders ... 115

U

UHF and VHF signal characteristics observed on a long knife-edge diffraction path ... 437
URSI, almost fifty years ... 317
Useful radiation from an underground antenna ... 89
Utlaut, W. E., Effect of antenna radiation angles upon HF radio signals propagated over long distances ... 167

V

Variance classification system for DF bearings, Brooke ... 255
Variances of position lines from fixes with unknown target positions, estimation ... 263
Very low frequencies (VLF) during a sudden ionospheric disturbance (SID), attenuation coefficients for propagation ... 543
VHF and UHF signal characteristics observed on a long knife-edge diffraction path ... 437
VLF and ELF radio waves by a horizontal magnetic dipole, excitation ... 305
VLF atmospheric noise, effect of receiver bandwidth on the amplitude distribution ... 299
VLF navigational aids, effects of the ionosphere ... 575
VLF propagation, a new approach to the mode theory ... 37
VLF propagation, comparison between mode theory and ray theory ... 357
VLF propagation perpendicular to the magnetic meridian, reflection from a sharply bounded ionosphere ... 455
VLF standard frequency and time signal broadcasting worldwide ... 617
Vogler, L. E., Smooth earth diffraction calculations for horizontal polarization ... 397
—, *C. F. Pappas, P. L. Rice*, Graphical determination of radio ray bending in an exponential atmosphere ... 175
Volland, H., Comparison between mode theory and ray theory of VLF propagation ... 357

W

Wait, J. R., A new approach to the mode theory of VLF propagation ... 37
—, A note concerning the excitation of ELF electromagnetic waves ... 481
—, On the theory of mixed-path ground-wave propagation on a spherical earth ... 401
—, *A. G. Jean, A. C. Murphy, D. F. Wasmundt*, Observed attenuation rate of ELF radio waves ... 475
Wasmundt, D. F., A. G. Jean, A. C. Murphy, J. R. Wait, Observed attenuation rate of ELF radio waves ... 475
Watt, A. D., R. W. Plush, W. W. Brown, A. H. Morgan, Worldwide VLF standard frequency and time signal broadcasting ... 617
Waveguides, characteristics for long-distance transmission ... 75
Waves, note concerning the excitation of ELF electromagnetic ... 481
Weather and reception level on a troposphere link—annual and short-term correlations ... 155
Wheeler, H. A., Radio-wave propagation in the earth's crust ... 189
—, Useful radiation from an underground antenna ... 89
Whistler ray paths, computation ... 485
Wilson, A. C., Measurements of low-angle radiation from a monopole ... 641
—, *H. V. Cottony*, A high-resolution rapid-scan antenna ... 101
Wilson, L. D., L. H. Heisler, Design of panoramic ionospheric recorders ... 629
Wind, solar ... 537
Worldwide VLF standard frequency and time signal broadcasting ... 617

Y

Yabroff, I., Computation of whistler ray paths ... 485

U.S. GOVERNMENT PRINTING OFFICE: 1961

Notice

The papers listed below will appear in early issues of RADIO PROPAGATION:

A survey of the very wide band and frequency independent antennas—1945 to the present. John Dyson.

Numerical investigation of the equivalent impedance of a wire grid parallel to the interface between two media. Tove Larsen.

Current on and input impedance of a cylindrical antenna. Yung Ming Chen and Joseph B. Keller.

Radar corner reflectors for linear or circular polarization. G. Latmiral and A. Sposito.

On the theory of wave propagation through a concentrically stratified troposphere with a smooth profile. Part II. Expansions of the rigorous solution. H. Bremmer.

On the propagation of VLF and ELF radio waves when the ionosphere is not sharply bounded. James R. Wait.

The fields of electric dipoles in sea water—the earth-atmosphere-ionosphere problem. Wallace L. Anderson.

The reflection of electromagnetic waves from thin ionized gaseous layers. F. H. Northover.

Reflection and transmission of radio waves at a continuously stratified plasma with arbitrary magnetic induction. J. Ralph Johler and John D. Harper, Jr.

On the diffraction of spherical radio waves by a finitely conducting spherical earth. L. C. Walters and J. R. Johler.

An approximate full wave solution for low frequency electromagnetic waves in an unbounded magneto-ionic medium. William C. Hoffman.

VHF radio propagation data for the Cedar Rapids-Sterling, Anchorage-Barrow, and Fargo-Churchill test paths, April 1951 through June 1958. (Digest of Tech. Note 79.) George R. Sugar and Kenneth W. Sullivan.

UNITED STATES
GOVERNMENT PRINTING OFFICE
DIVISION OF PUBLIC DOCUMENTS
WASHINGTON 25, D.C.

OFFICIAL BUSINESS

PENALTY FOR PRIVATE USE TO AVOID
PAYMENT OF POSTAGE, $300
(GPO)

JOURNAL OF RESEARCH of the National Bureau of Standards Vol. 65D, No. 6 November–December 1961
D. Radio Propagation

Contents

	Page
The solar wind. E. N. Parker.	537
Attenuation coefficients for propagation at very low frequencies (VLF) during a sudden ionospheric disturbance (SID). E. T. Pierce.	543
Dipole radiation in a conducting half space. R. K. Moore and W. E. Blair.	547
Reliability of atmospheric radio noise predictions. John R. Herman.	565
Effects of the ionosphere on VLF navigational aids. W. T. Blackband.	575
On the spectrum of terrestrial radio noise at extremely low frequencies. H. R. Raemer.	581
The nonsingular embedding of transition processes within a more general framework of coupled variables. John Heading.	595
Worldwide VLF standard frequency and time signal broadcasting. A. D. Watt, R. W. Plush, W. W. Brown, and A. H. Morgan.	617
Design of panoramic ionospheric recorders. L. H. Heisler and L. D. Wilson.	629
A quick method for estimating the stage of the sunspot cycle. W. B. Chadwick.	637
Measurements of low-angle radiation from a monopole. Alvin C. Wilson.	641
Publications of the Staff of the National Bureau of Standards.	647
Index to volume 65D. Radio Propagation, January to December 1961.	651

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. Price 70 cents (single copy)
Subscription price: $4.00 a year; 75 cents additional for foreign mailing

UNITED STATES GOVERNMENT PRINTING OFFICE, WASHINGTON : 1961